Phylogeography, Evolutionary History and Genetic Diversity of Sea Stars of the Genus *Astropecten* and Genetic Structure within the Atlanto-Mediterranean Species *A. aranciacus*

Dissertation

zur

Erlangung der naturwissenschaftlichen Doktorwürde

(Dr. Sc. Nat.)

vorgelegt der

Mathematisch-naturwissenschaftlichen Fakultät

der

Universität Zürich

von

Deborah E. Zulliger

von Madiswil BE

Promotionskomitee

Prof. Dr. Tony Wilson (Vorsitz)

Prof. Dr. Georg Ribi

Prof. Dr. Wolf Blankenhorn

Prof. Dr. Paul Ward

Zürich, 2009

CONTENTS

SUMMARY

The present study addresses issues related to the genetic diversity and evolutionary history of sea stars of the species-rich genus *Astropecten* (Asteroidea:Paxillosida:Astropectinidae) by inferring phylogenetic relationships among species on a global scale and investigating population structure at the species-level. Molecular markers were applied to resolve questions of taxonomy, diversity and speciation within *Astropecten* and to determine the population genetic structure within the Atlanto-Mediterranean species *Astropecten aranciacus*. Phylogenies were inferred using sequences of the mitochondrial DNA (mtDNA) and were dated using fossil records and geological events in order to reconstruct the evolutionary history of this genus. The population genetic structure of *A. aranciacus* was investigated by comparing results derived from mtDNA sequences with those from nuclear microsatellite markers. These microsatellite markers were novel and were therefore first tested and characterized. Results of this study resolve questions of phylogeography, speciation patterns and taxonomy and point out further areas in need to be investigated. This thesis also provides new nuclear molecular markers for asteroids and answers questions related to the genetic structure of marine invertebrate species with long-lived planktonic larval stages in the Atlanto-Mediterranean region.

The results of this study are structured into four chapters of which Chapters III + IV have been published in international journals. Chapter I has been submitted for peer-revision and Chapter II is in preparation for submission.

Chapter I

The aim of this study was to investigate speciation patterns and potential taxonomical inconsistencies within the genus *Astropecten* on a global scale using molecular markers. In addition, we assessed the possibility of morphological convergence between similar species found in different geographical regions and compared our results to the phylogeny proposed by Döderlein (1917), which was based on morphological characters. Given the taxonomical issues related to the "species-complex" of *A. irregularis,* a particular focus was set on species of the Mediterranean region. In total, we inferred phylogenetic relationships of 40 species of *Astropecten* collected worldwide using sequences of three mitochondrial DNA regions: 12S rRNA, 16S rRNA and COI. The resulting tree topologies were almost identical, with three major clades grouping species of the same geographical region together. These regions were: 1. Indo-Pacific; 2. east and west American coast; and 3. east Atlantic and Mediterranean. The phylogenetic affinity of the Mediterranean and east Atlantic species *A. aranciacus*, however, was not confidently resolved. Results also revealed species-complexes in taxa with a presumed large distribution area such as *A. polyacanthus* and *A. indicus*, suggesting the possibility of cryptic speciation within these complexes. Moreover, our results indicated morphological convergence in *Astropecten*, resolved several taxonomical issues and highlighted the necessity of an in-depth revision of this genus.

Chapter II

In taxa with planktotrophic larval stages, speciation events are presumably rare due to high effective population sizes and abundant gene flow. Despite the high potential for dispersal, a remarkable level of species diversity has evolved in the genus *Astropecten*. This raises questions about the evolutionary history of this group, the study of which is generally seen as a window to past speciation patterns and processes, and which may allow the determination of the origin of speciation. The objective of the study presented in this chapter was therefore to reconstruct the evolutionary history of *Astropecten* lineages, using geological events and fossil records to date a molecular phylogeny. Using the mtDNA sequence data produced and presented in Chapter I and after removing redundant taxa, I inferred a molecular phylogeny and calibrated the resulting tree using a) geological events only, b) fossil records only and c) a combination of geological events and fossil records. Using a relaxed clock method as implemented in the software MULTIDIVTIME I estimated divergence times for extant *Astropecten* lineages and compared the results from the different calibration categories. According to my results, recent lineages began to diverge in the Mid-Miocene, whereas the species diversity in the East Pacific and West Atlantic were most likely enhanced by the rise of the Panama Isthmus in the Late Miocene and Early Pliocene. Furthermore, the majority of exclusively Mediterranean species most likely evolved after the Messinian Salinity Crisis (Late Miocene, ~5 mya), suggesting a rather recent Mediterranean origin of these lineages. Ancient lineages, dating back to the Mid-Miocene, can be found in species from the South Pacific and in *A. aranciacus* of the Mediterranean and East Atlantic. Phylogenetic relationships of these lineages, however, could not be satisfactorily resolved with the mtDNA regions used in this study.

Chapter III

Due to its large body size and its historical abundance, *A. aranciacus* most likely plays an important role as a benthic predator throughout its distribution range in the Mediterranean and East Atlantic. However, the numbers of *A. aranciacus* have drastically declined in several locations within the Mediterranean during the last 20 years, raising questions of genetic diversity and population structure within this species. In order to investigate these issues, we employed polymorphic nuclear markers, such as microsatellite loci, which have proven to be useful tools when investigating the recent population genetic structure. A review of the existing literature revealed that the characterization of suitable microsatellite loci for echinoderms - and particularly for asteroids - has not been very extensive so far. Consequently, we tested and characterized microsatellite primers for markers which could be suitable to investigate the population genetic structure of *A. aranciacus*. For this purpose, an enriched microsatellite library was developed and positive clones were sequenced by Ecogenics GmbH (Zurich, Switzerland). We then designed and tested primers for several microsatellite inserts on *A. aranciacus*. Nine polymorphic microsatellite loci were identified and characterized using two populations of *A. aranciacus* and were cross-amplified with related asteroid species.

Chapter IV

The objective of this section was to investigate the impact of potential marine barriers on gene flow in high dispersal marine invertebrates by assessing the population genetic structure of the Atlanto-Mediterranean sea star *Astropecten aranciacus*. This species has a planktonic larval stage of up to 60 days and is therefore expected to have a high potential for dispersal and display less genetic structure than species without an extended planktonic stage. However, marine dispersal barriers are not always easy to identify and might constrain gene flow, leading to genetic population structure. For instance, the role of the Atlantic-Mediterranean division or the Siculo-Tunisian Strait as potential barriers to gene flow have increasingly been investigated for various planktotrophic invertebrate species. In order to investigate potential marine barriers to gene flow in *A. aranciacus*, we applied molecular genetic methods using both mitochondrial and nuclear markers in samples collected throughout the Mediterranean and the contiguous East Atlantic region. The control region of the mitochondrial DNA was sequenced and nuclear data from microsatellite loci were produced to obtain a complementary picture of the genetic diversity in *A. aranciacus*. Results deriving from both markers were coherent and showed similar patterns of genetic structure in *A. aranciacus*. Microsatellite and mtDNA data sets demonstrated a clear pattern of isolation-by-distance and revealed a significant genetic differentiation of the population from the island of Madeira in the Atlantic to most other populations. Microsatellite loci further indicated possible genetic differentiation between the East Atlantic, the West and the East Mediterranean basins.

ZUSAMMENFASSUNG

Die vorgelegte Arbeit behandelt Fragen im Zusammenhang mit der genetischen Vielfalt, Taxonomie, Artbildung und Evolutionsgeschichte von Seesternen der Gattung *Astropecten* (Asteroidea:Paxillosida:Astropectinidae). Dazu wurden einerseits auf globaler Ebene die Verwandtschaften zwischen den Arten dieser Gruppe und andererseits die Populationsstruktur innerhalb der Art *Astropecten aranciacus*, welche im Ostatlantik und Mittelmeer vorkommt, ermittelt. Dies geschah unter Verwendung molekularer Marker. Die Phylogenien wurden anhand von Sequenzendaten der mitochondrialen DNA (mtDNA) abgeleitet. Anschliessend wurden diese mit Hilfe von Fossilien und geologischen Ereignissen datiert, um die Evolutionsgeschichte dieser Gattung zu rekonstruieren. Die Populationsstruktur von *A. aranciacus* wurde untersucht, indem Ergebnisse von mtDNA Sequenzen mit denjenigen von Mikrosatelliten verglichen wurden. Diese Mikrosatelliten waren neuartig und wurden deshalb vorher getestet und beschrieben. Die Ergebnisse dieser Studien klären Fragen zur Phylogeografie und Taxonomie, geben Hinweise auf Artbildungsmuster und zeigen weitere interessante Gebiete auf, die zu untersuchen wären. Zudem wurden durch diese Arbeit neue nukleare molekulare Marker für Asteroida zur Verfügung gestellt und Fragen zur genetischen Struktur von marinen Wirbellosen mit langlebigen planktonischen Larvalstadien beantwortet.

Die Ergebnisse dieser Arbeit sind in vier Kapitel gegliedert, wovon Kapitel III und IV in internationalen Zeitschriften publiziert wurden. Kapitel I wurde zur „peer-revision“ eingereicht, und Kapitel II wird zur Publikation vorbereitet.

Kapitel I

Das Ziel dieser Arbeit war es, Artbildungsmuster und potentielle taxonomische Ungereimtheiten innerhalb der Gattung *Astropecten* auf globaler Ebene unter Anwendung von molekularen Markern zu ermitteln. Zudem wurde beurteilt, ob in dieser Gattung morphologische Konvergenz zwischen ähnlichen Arten in verschiedenen geographischen Gebieten vorkommt. Die Ergebnisse wurden mit dem Stammbaum von Döderlein (1917) verglichen, welcher aufgrund von morphologischen Kriterien erstellt wurde. Da taxonomische Unklarheiten innerhalb des Artenkomplexes von *A. irregularis* bestehen, einer Gruppe, die im Mittelmeer und Ostatlantik vorkommt, wurde diese Region noch speziell behandelt. Die Phylogenie wurde insgesamt mit 40 *Astropecten*-Arten rekonstruiert, anhand drei mitochondrialer DNA-Regionen: 12S rRNA, 16S rRNA und COI. Die Topologien der resultierenden Phylogenien waren fast identisch und wiesen drei Hauptstämme auf, welche jeweils Arten der gleichen geographischen Region zusammengruppierten. Diese Regionen waren: 1. Indo-West Pazifik; 2. Ost- und Westküste von Amerika; und 3. Ostatlantik und Mittelmeer. Die phylogenetische Affinität von *A. aranciacus*, einer Art, die im Mittelmeer und im Ostatlantik vorkommt, wurde jedoch nicht befriedigend aufgelöst. Die Ergebnisse zeigten auch auf, dass Taxa mit einem mutmasslich grossen Verbreitungsgebiet wie *A. polyacanthus* and *A. indicus* eigentlich

Artenkomplexe sind. Dies deutet darauf hin, dass kryptische Artbildung innerhalb dieser Komplexe möglich ist. Die Resultate weisen zudem auf das Vorkommen von morphologischer Konvergenz bei *Astropecten* hin, klären verschiedene taxonomische Fragen und zeigen die Notwendigkeit einer tiefgehenden Revision dieser Gattung auf.

Kapitel II

Bei Taxa, welche planktotrophische Larvalstadien aufweisen, wird oft angenommen, dass Artbildungsereignisse aufgrund der grossen Populationsgrössen und des reichlichen Genflusses eher selten sind. Obwohl Seesterne der Gattung *Astropecten* planktotrophische Larvalstadien aufweisen und daher vermutlich ein hohes Verbreitungspotential haben, ist in dieser Gruppe im Vergleich zu anderen Echinodermen eine bemerkenswerte Artenvielfalt entstanden. Dies wirft die Frage nach der Evolutionsgeschichte dieser Gruppe auf. Durch die Rekonstruktion der Evolutionsgeschichte einer Gruppe kann ein Einblick in Artbildungsmuster und –prozesse gewonnen werden und es können allenfalls auch Hinweise zum Ursprungsort der Artbildung geliefert werden. Das Ziel dieser Studie war daher, die Evolutionsgeschichte der Stammlinien von *Astropecten* zu ermitteln, indem eine molekulare Phylogenie mittels geologischen Ereignissen und Fossilien datiert wurde. Die mtDNA Sequenzdaten, welche bereits in Kapitel I produziert und präsentiert worden sind, wurden dazu erneut verwendet. Eine molekulare Phylogenie wurde rekonstruiert, und der daraus resultierende Stammbaum wurde kalibriert mittels a) geologischen Ereignissen, b) Fossilien und c) einer Kombination von geologischen Ereignissen und Fossilien. Die Methode der „entspannten Uhr“, wie sie in der Software MULTIDIVTIME implementiert ist, wurde benutzt, um den Zeitpunkt der Divergenz der bestehen Stammlinien von *Astropecten* zu berechnen. Die Ergebnisse der verschiedenen Kalibrierungen wurden anschliessend verglichen. Gemäss den Ergebnissen dieser Studie begannen rezente Stammlinien sich im mittleren Miozän abzuspalten, wobei die Artenvielfalt im Ostpazifik und Westatlantik vermutlich durch den Anstieg des Isthmus von Panama im späten Miozän und frühen Pliozän gefördert wurde. Zudem sind die meisten Arten, welche ausschliesslich das Mittelmeer bewohnen, nach der Messinischen Salinitätskrise (spates Miozän, frühes Pliozän ~5 mya) entstanden, was auf einen relativ jungen Ursprung dieser Stammlinien im Mittelmeer hindeutet. Zu den älteren Stammlinien, welche ins mittlere Miozän zurückdatieren, gehören Arten des Südpazifiks und *A. aranciacus*, eine Art, welche im Ostatlantik und Mittelmeer verbreitet ist. Die Verwandtschaft dieser älteren Stammlinien konnte jedoch mit den in dieser Studie verwendeten mtDNA Regionen nicht zufrieden stellend geklärt werden.

Kapitel III

Die Art *A. aranciacus* wird innerhalb ihres Verbreitungsgebietes im Ostatlantik und Mittelmeer aufgrund ihrer erheblichen Körpergrösse und ihres (ehemals) häufigen Vorkommens als ökologisch bedeutender benthischer Räuber betrachtet. Die Anzahl Individuen dieser Art hat jedoch in verschiedenen Regionen des Mittelmeeres innerhalb der letzten 20 Jahre erheblich abgenommen. Die Frage nach der genetischen Vielfalt

und Populationsstruktur innerhalb dieser Art liegt daher nahe. Um diese Frage zu beantworten, wurden polymorphe Mikrosatelliten verwendet, welche sich als nützliches Werkzeug zur Ermittlung der populationsgenetischen Struktur bewährt haben. Eine Prüfung der bestehenden Literatur hat jedoch gezeigt, dass die Charakterisierung von geeigneten Mikrosatelliten-Loci für Echinodermen – und insbesondere für Asteroida – nicht besonders ergiebig ist. Aufgrund dessen wurden in dieser Arbeit Primer für verschiedene Mikrosatelliteninserte in *A. aranciacus* entwickelt und getestet, anhand der positiven Klone einer angereicherten Mikrosatelliten-Bibliothek, welche die Firma Ecogenics GmbH (Zürich, Schweiz) erstellt hat. Neun polymorphe Mikrosatelliten-Loci in *A. aranciacus* wurden identifiziert und anhand von zwei Populationen beschrieben. Anschliessend wurde die Nützlichkeit dieser Primer auch in verwandten Seesternarten getestet.

Kapitel IV

In dieser Studie wurde die Frage behandelt, ob der Genfluss bei marinen Wirbellosen mit einem hohen Verbreitungspotential durch eventuelle marine Barrieren beeinflusst wird. Zu diesem Zweck wurde die Populationsstruktur von *Astropecten aranciacus*, einer im Ostatlantik und Mittelmeer vorkommenden Seesternart, untersucht. Da das planktische Larvalstadium dieser Art bis zu 60 Tagen dauert, kann angenommen werden, dass sie ein hohes Verbreitungspotenzial hat und eine tiefere genetische Struktur aufweist als Arten ohne entsprechend lange planktische Phase. Marine Verbreitungsbarrieren, welche oft nicht einfach zu erkennen sind, können jedoch den Genfluss behindern, was zu einer deutlichen Populationsstruktur führen kann. Ob der Trennungsbereich zwischen Mittelmeer und Atlantik oder die "sizilianisch-tunesische Strasse" Barrieren für den Genfluss darstellen, wurde bereits für diverse Wirbellose mit planktischen Larvalstadien untersucht. Um die Frage, ob der Genfluss in *A. aranciacus* durch marine Barrieren behindert wird, zu beantworten, haben wir molekulargenetische Methoden angewendet. Dabei wurden sowohl mitochondriale als auch nukleare Marker in Individuen untersucht, welche aus verschiedenen Populationen des Mittelmeeres und des angrenzenden Ostatlantiks stammten. Die Kontrollregion der mitochondrialen DNA (mtDNA) wurde sequenziert und nukleare Daten von Mikrosatelliten-Loci produziert, um ein umfassendes Bild der genetischen Vielfalt bei *A. aranciacus* zu erhalten. Die Ergebnisse der beiden Markertypen waren kohärent und ergaben eine ähnliche genetische Populationsstruktur von *A. aranciacus*. Sowohl Mikrosatelliten-Loci als auch mtDNA wiesen ein klares Muster einer Isolation durch Entfernung auf und enthüllten eine signifikante genetische Differenzierung der Population von Madeira im Ostatlantik zu den anderen Populationen. Mikrosatelliten-Loci wiesen zudem auf eine mögliche genetische Differenzierung zwischen Ostatlantik, westlichem und östlichem Mittelmeer hin.

GENERAL INTRODUCTION

Marine invertebrates with long-lived planktotrophic larvae are thought to rarely undergo speciation due to their capacity for dispersal and large effective population sizes (Palumbi 1992; Palumbi 1994; Riginos and Victor 2001). Thus, there are few genera with high species diversity for example among echinoderms with planktonic dispersal. However, this trend is not universal among all echinoderms, as for example in sea stars of the genus *Astropecten* (Asteroidea:Paxillosida:Astropectinidae), for which over 150 species are currently discribed (Say 1825; Gray 1841; Müller and Troschel 1842; Perrier 1875/6; Agassiz 1877; Sladen 1889; Ludwig 1897; Fisher 1906; Koehler 1909; Koehler 1910; Fisher 1911; Fisher 1913; Verrill 1914; Verrill 1915; Döderlein 1917; Koehler 1924; Clark and Downey 1992). This is remarkable, because most species of *Astropecten* presumably have long planktonic larval stages, which in the case of e.g., *A. aranciacus* can last up to 60 days (Hörstadius 1938). To explain this discrepancy, either past or current barriers to dispersal must have been present. Alternatively, the larval stage in *Astropecten* species may not always be as long-lived as assumed due to biological or environmental factors.

Reconstructing phylogenetic relationships can help to determine speciation patterns and processes, such as e.g., sympatric versus allopatric speciation. Several studies have used molecular markers in echinoderms with planktotrophic larvae to infer relationships on a global scale in order to address such questions (e.g., Lessios et al. 1999; Williams 2000; Lessios et al. 2001; Lessios et al. 2003; Waters et al. 2004). However, these studies have only included genera with few, mostly allopatric species and therefore do not answer questions concerning e.g., morphological convergence in echinoderms.

Molecular markers such as DNA sequence data, are not only useful tools for the investigation of speciation patterns and processes, but can also be used to detect cryptic species and morphological variations within the same species. As species of sea stars have typically been described on the basis of morphological characteristics, the high variability in the genus *Astropecten* has resulted in the classification of several subspecies, variations and local forms. Moreover, many new species have been described based on juvenile or poorly preserved specimens, and thus, the validity of several species is questionable, suggesting that a revision of this genus is required using modern techniques as in molecular genetics.

In *Astropecten,* phylogenetic relationships have not yet been inferred either by a cladistic approach or on the basis of molecular markers. However, in 1917, Döderlein published a monograph on the genus *Astropecten* and its evolutionary history based on morphological characters. Döderlein proposed that in some cases morphologically similar species living in similar habitats but in different geographical regions may be closely related. Until now, this assumption has not yet been tested in a phylogenetic framework by reconstructing a phylogeny based on e.g., DNA sequence data.

The high species-diversity in a marine invertebrate with a large dispersal capacity also raises questions related to the evolutionary history of the group as well as the timing of the presumed species radiation. Determining the age of lineage divergence within taxa is essential when studying evolutionary patterns, speciation events and rates of evolution (Smith and Peterson 2002). In order to estimate divergence times of

lineages, phylogenetic trees, e.g., inferred from molecular sequence data, can be dated using fossils or geological events. The origin of numerous taxa has been investigated this way, including marine invertebrates such as crustaceans (Perez-Losada et al. 2004; Perez-Losada et al. 2008; Tinn and Oakley 2008), mollusks (Wilke et al. 2000; Wood et al. 2007; Frey and Vermeij 2008) and echinoids (Smith et al. 2006). However, to the best of my knowledge, no studies until now have estimated divergence times in asteroids.

As mentioned earlier, gene flow in marine species with long planktonic larval stages may be more restricted than generally assumed. For instance, the role of the Atlantic-Mediterranean division as a potential barrier to gene flow in various planktotrophic invertebrate species has increasingly been investigated (e.g., (Borsa et al. 1997; Launey et al. 2002; Diaz-Almela et al. 2004; Duran et al. 2004; Stamatis et al. 2004; Saavedra and Pena 2005; Stamatis et al. 2006; Calderon et al. 2008). Other barriers to gene flow within the Mediterranean may also exist in the form of an east-west divide at the Siculo-Tunisian Strait and/or hydrogeographic isolation of the Aegean, Ionian and Adriatic Seas (Perez-Losada et al. 2007).

Many studies have used genetic tools such as mitochondrial DNA (mtDNA), nuclear DNA or a combination of the two to analyze genetic structure in high dispersal Atlanto-Mediterranean invertebrates (e.g., Féral et al. 1995; Zane et al. 2000; Launey et al. 2002; Diaz-Almela et al. 2004; Duran et al. 2004; Roman and Palumbi 2004; Stamatis et al. 2004; Triantafyllidis et al. 2005; Peijnenburg et al. 2006; Calderon et al. 2008). However, incongruent conclusions regarding the influence of the Atlantic-Mediterranean division on population structuring were drawn in these studies depending not only on the species investigated but also on the genetic markers used and the sampling pattern. For instance, only moderate genetic differentiation was revealed between Atlantic and Mediterranean populations of the sea urchin *Paracentrotus lividus* based on mtDNA sequences (Duran et al. 2004), whereas a sharp break was detected between the two basins when combining mitochondrial and nuclear markers and applying a more extensive sampling (Calderon et al. 2008).

Comparing results derived from both mitochondrial and nuclear data can provide a more comprehensive picture, but to our knowledge no study has yet been conducted assessing the population structure in sea stars with planktotrophic larvae in the Atlanto-Mediterranean region. Suitable molecular markers to investigate this aspect include the control region of the mitochondrial DNA, a region that has the highest rate of evolutionary change of any mtDNA region (Aquadro and Greenberg 1983; Parsons et al. 1997) and polymorphic nuclear microsatellite loci, which are believed to be neutral and have been shown to be more variable compared to e.g., allozyme data (Shaw et al. 1999; Estoup et al. 1998; Perez-Losada et al. 2002).

The objective of this thesis was to address issues related to phylogenetic relationships and evolutionary history in sea stars of the species-rich genus *Astropecten* (Asteroidea:Paxillosida) and to assess the genetic diversity at both an inter- and intra-species level using molecular markers.

In **Chapter I**, we inferred a molecular phylogeny using a global sampling of *Astropecten* species and mtDNA sequence data in order to investigate species diversity and speciation patterns in this genus and to assess the possibility of cryptic speciation and morphological convergence. Sampling of specimens, production of DNA sequences and data processing and analysis were my contribution to this chapter, whereas Harilaos Lessios helped to improve the manuscript for publication.

In **Chapter II** I then addressed questions related to the evolutionary history of *Astropecten*, such as the approximate divergence time of *Astropecten* lineages, speciation processes and the influence of geological events on lineage divergence. I used the sequence data produced in Chapter I and applied molecular dating techniques by calibrating nodes in the phylogeny using geological events and fossil records. A relaxed molecular clock method was then applied to estimate divergence times of extant *Astropecten* lineages.

To assess genetic diversity at the species-level and to determine potential barriers to larval dispersal in the Atlanto-Mediterranean realm, we attempted to investigate population structure within the high dispersal species *Astropecten aranciacus*. For this, in **Chapter III**, we first tested primers for potentially suitable microsatellite loci in *A. aranciacus*, of which we characterized nine and cross-amplified them to test their suitability in other asteroid species. My contribution to this chapter was to complete the primer testing, to analyze the data and to write the manuscript. Markus Ruch initiated the project and did preliminary testing of primers, while Samuel Tanner helped in the lab and Georg Ribi collected samples.

In **Chapter IV** we investigated the influence of potential marine barriers to gene flow in high dispersal marine invertebrates in the Atlanto-Mediterranean region by assessing the population genetic structure of *A. aranciacus*. We did this by selecting the four best microsatellite loci for genotyping from Chapter III and by sequencing DNA of the mitochondrial control region. Moreover, results derived from nuclear and mitochondrial markers were compared in order to obtain a comprehensive picture of the population structure and to assess potential method related differences. Samuel Tanner produced and analyzed mtDNA sequence data and contributed to the manuscript, while my participation consisted of genotyping microsatellite loci, analyzing the remaining data and writing up the manuscript for publication. Markus Ruch and Georg Ribi assisted with sampling and did preliminary conceptual work.

References

Agassiz A (1877) North American starfishes. Mem Mus Comp Zool **5**: 1-136.

Aquadro CF, Greenberg BD (1983) Human Mitochondrial-DNA Variation and Evolution - Analysis of Nucleotide-Sequences from 7 Individuals. Genetics **103**: 287-312.

Borsa P, Blanquer A, Berrebi P (1997) Zoogéographie intraspécifique de la mer Méditerranée. Analyse des données génétiques populationnelles sur seize espèces atlanto-méditerranéennes (Poissons et Invertèbres). Vie et Milieu **47**: 95-305.

Calderon I, Giribet G, Turon X (2008) Two markers and one history: phylogeography of the edible common sea urchin *Paracentrotus lividus* in the Lusitanian region. Mar Biol **154**: 137-151. doi: 10.1007/s00227-008-0908-0.

Clark AM, Downey ME (1992) Starfishes of the Atlantic. Chapman & Hall, London, pp 794.

Diaz-Almela E, Boudry P, Launey S, Bonhomme F, Lapegue S (2004) Reduced female gene flow in the European flat oyster *Ostrea edulis*. J Hered **95**: 510-516. doi: 10.1093/jhered/esh073.

Döderlein L (1917) Die Asteriden der Siboga-Expedition. I. Die Gattung *Astropecten* und ihre Stammesgeschichte. In: Brill EJ (ed) Siboga-Expeditie. Uitkomsten op zoölogisch, botanisch, ozeanographisch en geologisch gebied verzameld in Nederlandsch Oost-Indie 1899-1900 aan boord H.M. "Siboga". 46 (a), Leiden, pp 191.

Duran S, Palacin C, Becerro MA, Turon X, Giribet G (2004) Genetic diversity and population structure of the commercially harvested sea urchin *Paracentrotus lividus* (Echinodermata, Echinoidea). Mol Ecol **13**: 3317-3328. doi: 10.1111/j.1365-294X.2004.02338.x.

Estoup A, Rousset F, Michalakis Y, Cornuet JM, Adriamanga M, Guyomard R (1998) Comparative analysis of microsatellite and allozyme markers: a case study investigating microgeographic differentiation in brown trout (*Salmo trutta*). Mol Ecol **7**: 339–353. doi:10.1046/j.1365-294X.1998.00362.x.

Féral J-P, Poulin E, Derelle E, Gallardo S, Chambon C (1995) Genetic differentiation of *Echinocardium chordatum* as revealed by allozymes and RNA sequencing. In: Emson R, Smith A, Campbell A (eds) Echinoderm research 1995, Balkema, Rotterdam, pp 41-42.

Fisher WK (1906) New Starfishes from the Pacific Coast of North America. Proc Wash Acad Sci **8**: 11-139.

Fisher WK (1911) Asteroidea of the North Pacific and adjacent waters. Part 1. Phanerozonia and Spinulosa. Bull U.S. Nat Mus **76**, pp 420.

Fisher WK (1913) Four new genera and fifty-eight new species of starfishes from the Philippine Islands, Celebes, and the Moluccas. Proc U.S. Nat Mus **43**: 599-648.

Frey MA, Vermeij GJ (2008) Molecular phylogenies and historical biogeography of a circumtropical group of gastropods (Genus : Nerita): Implications for regional diversity patterns in the marine tropics. Mol Phylogenet Evol **48**: 1067-1086. doi: 10.1016/j.ympev.2008.05.009.

Gray JE (1841) A synopsis of the genera and species of the class Hypostoma (Asterias, Linnaeus). Ann Mag Nat Hist **1**: 175-184.

Hörstadius S (1938) Über die Entwicklung von *Astropecten aranciacus* L. Pubbl Stn Zool Napoli **17**: 221-312.

Koehler R (1909) Echinodermes provenant des campagnes du yacht Princesse-Alice. Resultats des campagnes scientifiques (Monaco) **34**: 1-317.

Koehler R (1910) Shallow-Water Asteroidea. Echinoderma of the Indian Museum, Calcutta, Indian Museum **191**, pp 206.

Koehler R (1924) Les Echinodermes des Mers D'Europe. Librarie Octave Doin, Paris pp 362.

Launey S, Ledu C, Boudry P, Bonhomme F, Naciri-Graven Y (2002) Geographic structure in the European flat oyster (*Ostrea edulis* L.) as revealed by microsatellite polymorphism. J Hered 93: 331-338. doi: 10.1093/jhered/93.5.331.

Lessios HA, Kane J, Robertson DR (2003) Phylogeography of the pantropical sea urchin Tripneustes: Contrasting patterns of population structure between oceans. Evolution **57**: 2026-2036.

Lessios HA, Kessing BD, Pearse JS (2001) Population structure and speciation in tropical seas: Global phylogeography of the sea urchin Diadema. Evolution **55**: 955-975.

Lessios HA, Kessing BD, Robertson DR, Paulay G (1999) Phylogeography of the pantropical sea urchin Eucidaris in relation to land barriers and ocean currents. Evolution **53**: 806-817.

Ludwig H (1897) Die Seesterne des Mittelmeeres. Fauna und Flora des Golfes von Neapel und der angrenzenden Meeres-Abschnitte. Zool Stn Neapel **24**, pp 491.

Müller J, Troschel FH (1842) System der Asteriden. Bieweg & Sohn, Braunschweig, pp 132.

Palumbi SR (1992) Marine Speciation on a Small Planet. Trends Ecol Evol **7**: 114-118.

Palumbi SR (1994) Genetic-Divergence, Reproductive Isolation, and Marine Speciation. Ann Rev Ecol Syst **25**: 547-572.

Parsons TJ, Muniec DS, Sullivan K, Woodyatt N, AllistonGreiner R, Wilson MR, Berry DL, Holland KA, Weedn VW, Gill P, Holland MM (1997) A high observed substitution rate in the human mitochondrial DNA control region. Nature Genetics **15**: 363-368. doi: 10.1038/ng0497-363.

Peijnenburg K, Fauvelot C, Breeuwer AJ, Menken SBJ (2006) Spatial and temporal genetic structure of the planktonic *Sagitta setosa* (Chaetognatha) in European seas as revealed by mitochondrial and nuclear DNA markers. Mol Ecol **15**: 3319-3338. doi: 10.1111/j.1365-294X.2006.03002.x.

Perez-Losada M, Guerra A, Carvalho GR, Sanjuan A, Shaw PW (2002) Extensive population subdivision of the cuttlefish *Sepia officinalis* (Mollusca : Cephalopoda) around the Iberian Peninsula indicated by microsatellite DNA variation. Heredity **89**: 417-424. doi: 10.1038/sj.hdy.6800160.

Perez-Losada M, Harp M, Hoeg JT, Achituv Y, Jones D, Watanabe H, Crandall KA (2008) The tempo and mode of barnacle evolution. Mol Phylogenet Evol **46**: 328-346. doi: 10.1016/j.ympev.2007.10.004.

Perez-Losada M, Hoeg JT, Crandall KA (2004) Unraveling the evolutionary radiation of the thoracican barnacles using molecular and morphological evidence: A comparison of several divergence time estimation approaches. Syst Biol **5**3: 244-264. doi: 10.1080/10635150490423458.

Perez-Losada M, Nolte MJ, Crandall KA, Shaw PW (2007) Testing hypotheses of population structuring in the Northeast Atlantic Ocean and Mediterranean Sea using the common cuttlefish *Sepia officinalis*. Molecular Ecology 16: 2667-2679. DOI: 10.1111/j.1365-294X.2007.03333.x

Perrier E (1875/6) Révision de la collection de Stellerides du Muséum d'Histoire Naturelle de Paris. Arch Zool Exp Gen **4**: 265-450.

Riginos C, Victor BC (2001) Larval spatial distributions and other early life-history characteristics predict genetic differentiation in eastern Pacific blennioid fishes. Proc Roy Soc B **268**: 1931-1936.

Roman J, Palumbi SR (2004) A global invader at home: population structure of the green crab, *Carcinus maenas*, in Europe. Mol Ecol **13**: 2891-2898. doi: 10.1111/j.1365-294X.2004.02255.x.

Saavedra C, Pena JB (2005) Nucleotide diversity and Pleistocene population expansion in Atlantic and Mediterranean scallops (*Pecten maximus* and *P-jacobaeus*) as revealed by the mitochondrial 16S ribosomal RNA gene. J Exp Mar Biol Ecol **323**: 138-150. doi: 10.1016/j.jembe.2005.03.006.

Say T (1825) On the species of the Linnean genus Asterias inhabiting the coast of the United States. J Acad Nat Sci Phil **5**: 151-154.

Shaw PW, Pierce GJ, Boyle PR (1999) Subtle population structuring within a highly vagile marine invertebrate, the veined squid *Loligo forbesi*, demonstrated with microsatellite DNA markers. Molecular Ecology **8**: 407-417. doi: 10.1046/j.1365-294X.1999.00588.x.

Sladen WP (1889) Report on the Asteroidea. Report on the scientific results of the voyage of the H. M. S. Challenger during the years 1873-1876. Zoology **30**: 1-893.

Smith AB, Peterson KJ (2002) Dating the time of origin of major clades: Molecular clocks and the fossil record. Ann Rev Earth Planet Sci **30**: 65-88.

Smith AB, Pisani D, Mackenzie-Dodds JA, Stockley B, Webster BL, Littlewood TJ (2006) Testing the molecular clock: Molecular and paleontological estimates of divergence times in the echinoidea (Echinodermata). Mol Biol Evol **23**: 1832-1851. doi: 10.1093/molbev/msl039.

Stamatis C, Triantafyllidis A, Moutou KA, Mamuris Z (2004) Mitochondrial DNA variation in northeast atlantic and mediterranean populations of norway lobster, *Nephrops norvegicus*. Mol Ecol **13**: 1377-1390. doi: 10.1111/j.1365-294X.2004.02165.x.

Stamatis C, Triantafyllidis A, Moutou KA, Mamuris Z (2006) Allozymic variation in Northeast Atlantic and Mediterranean populations of Norway lobster, *Nephrops norvegicus*. Mar Sci **63**: 875-882. doi: 10.1016/j.icesjms.2006.01.006.

Tinn O, Oakley TH (2008) Erratic rates of molecular evolution and incongruence of fossil and molecular divergence time estimates in Ostracoda (Crustacea). Mol Phylogenet Evol **48**: 157-167. doi: 10.1016/j.ympev.2008.03.001.

Triantafyllidis A, Apostolidis AP, Katsares V, Kelly E, Mercer J, Hughes M, Jorstad K, Tsolou A, Hynes R, Triantaphyllidis C (2005) Mitochondrial DNA variation in the European lobster (*Homarus gammarus*) throughout the range. Mar Biol **146**: 223-235. doi: 10.1007/s00227-004-1435-2.

Verrill AE (1914) Monograph of the shallow-water Starfishes of the North Pacific Coast from the Arctic Ocean to California. Harriman Alaska Series **14**: 1-420.

Verrill AE (1915) Report on the Starfishes of the West Indies, Florida, and Brazil. Bull Lab Nat Hist **7**: 2-232.

Waters JM, O'Loughlin PM, Roy MS (2004) Molecular systematics of some Indo-Pacific asterinids (Echinodermata, Asteroidea): does taxonomy reflect phylogeny? Mol Phylogenet Evol **30**: 872-878. doi: 10.1016/j.ympev.2003.08.019.

Wilke T, Rolan E, Davis GM (2000) The mudsnail genus Hydrobia s.s. in the northern Atlantic and western Mediterranean: a phylogenetic hypothesis. Mar Biol **137**: 827-833.

Williams ST (2000) Species boundaries in the starfish genus Linckia. Mar Biol **136**: 137-148.

Wood AR, Apte S, MacAvoy ES, Gardner JPA (2007) A molecular phylogeny of the marine mussel genus Perna (Bivalvia : Mytilidae) based on nuclear (ITS1&2) and mitochondrial (COI) DNA sequences. Mol Phylogenet Evol **44**: 685-698. doi: 10.1016/j.ympev.2006.12.019.

Zane L, Ostellari L, Maccatrozzo L, Bargelloni L, Cuzin-Roudy J, Buchholz F, Patarnello T (2000) Genetic differentiation in a pelagic crustacean (*Meganyctiphanes norvegica* : Euphausiacea) from the North East Atlantic and the Mediterranean Sea. Mar Biol **136**: 191-199. doi: 10.1007/s002270050676.

CHAPTER I

Phylogenetic relationships in the genus *Astropecten* (Asteroidea:Paxillosida:Astropectinidae) on a global scale: molecular evidence for morphological convergence, species-complexes and possible cryptic speciation

Deborah E. Zulliger, Harilaos A. Lessios

Abstract

With over 150 described species, *Astropecten* (Asteroidea:Paxillosida:Astropectinidae) is one of the most species-rich genera among sea stars. This diversity is remarkable, because most species of *Astropecten* have a long-lived planktotrophic larval stage, which would be expected to lead to a low speciation rate. The taxonomy of this genus is complex and not well resolved, and phylogenetic relationships have only been addressed in the beginning of the last century. In order to resolve general taxonomic issues, identify speciation patterns and estimate species diversity within the genus *Astropecten,* we inferred a molecular phylogeny of 118 specimens of *Astropecten* belonging to 40 species from around the world using mitochondrial DNA (mtDNA) sequences of 12S rRNA, 16S rRNA and cytochrome oxidase subunit I (COI). We compared the resulting molecular phylogeny to a previously published morphological one by Döderlein and investigated the possibility of morphological convergence in species from different geographic regions. We also tried to identify potentially problematic descriptions and/or signs of cryptic speciation in *Astropecten.* Finally, by sampling all known species of the Mediterranean, we addressed taxonomic issues in species that inhabit this sea. The resulting global molecular phylogeny exhibited three main clades, each containing specimens of the same geographic region. The three geographic regions are: 1. The Indo-Pacific; 2. The Neotropics; and 3. The eastern Atlantic and Mediterranean. Phylogenetic inferences based on mtDNA indicate that morphological and ecological convergence has taken place in *Astropecten,* resulting in allopatric non-sister taxa with similar morphologies and habitat preferences. The comparison to Döderlein's morphological phylogeny reveals congruence on the whole, but many discrepancies on a local scale, indicating that meaningful morphological characters are not easily identified and categorized in *Astropecten.* Our results also indicate that *A. polyacanthus* and *A. indicus* are species-complexes*;* cryptic speciation may have occurred within each of these morphospecies. Furthermore, many variants, previously presumed to be conspecific, exhibit genetic distances large enough to justify recognizing them as separate species. For instance, the data suggest that *A. irregularis* var. *pentacanthus* in the Mediterranean can be considered a species of its own and be referred to as *A. pentacanthus*.

Introduction

Marine invertebrates with long-lived planktotrophic larvae are thought to rarely undergo speciation due to their enormous capacity for dispersal and large effective population sizes (Palumbi 1992; Palumbi 1994; Riginos and Victor 2001). Thus, there are few genera with high species diversity among echinoderms. However, this trend is not universal among all echinoderms, and therefore, estimating probability of speciation from the duration of the larval stage alone can be misleading. This has been shown in several studies using genetic markers, which have found that levels of gene flow and extent of species range cannot always be predicted from larval duration (e.g., Sponer and Roy 2002; Waters et al. 2004; Paulay and Meyer 2006; Wilson et al. 2007).

Genus *Astropecten* (Asteroidea:Paxillosida): taxonomy

An exception to the pattern of low species diversity in echinoderms with mostly long-lived planktonic larval stages is the sea star genus *Astropecten* (Asteroidea: Paxillosida: Astropectinidae) with approximately 150 species described to date (Say 1825; Gray 1841; Müller and Troschel 1842; Perrier 1875/6; Agassiz 1877; Sladen 1889; Ludwig 1897; Fisher 1906; Koehler 1909; Koehler 1910; Fisher 1911; Fisher 1913; Verrill 1914; Verrill 1915; Döderlein 1917; Koehler 1924; Clark and Downey 1992). As species of sea stars typically remain defined by morphological characteristics, the high phenotypic variability in *Astropecten* has resulted in the designation of several subspecies, variations and local forms. Many new species have been described based on juvenile or badly preserved specimens, such as *A. exiguus* Ludwig, *A. hermatophilus* Sladen, *A. ibericus* Perrier, *A. progressor* Döderlein and *A. spiniphorus* Madsen. The validity of several of these species is questionable, and extensive revision of this genus is required using modern techniques.

In the Mediterranean Sea, the occurrence of six *Astropecten* species has been broadly accepted (e.g., Döderlein 1917; Koehler 1924; Tortonese 1965; Clark and Downey 1992). These species are: *A. aranciacus* Linnaeus, *A. bispinosus* Otto, *A. irregularis* Pennant, *A. jonstoni* Delle Chiaje, *A. platyacantus* Philippi and *A. spinulosus* Philippi. Most authors regard *A. aranciacus*, *A. bispinosus*, *A. jonstoni*, *A. platyacanthus* and *A. spinulosus* as clearly defined species; however, Koehler (1924) and Ludwig (1897) considered *A. platyacanthus* as a variant of *A. bispinosus*. In *A. irregularis*, several subspecies have been described (Döderlein 1917; Koehler 1924; Tortonese 1965; Clark and Downey 1992): the typical Mediterranean form *A. irregularis pentacanthus* Delle Chiaje, the South African form *A. irregularis pontoporaeus* Sladen and the East Atlantic form *A. irregularis irregularis* Pennant, occurring from Norway to Cape Verde. Some authors consider *A. irregularis irregularis* and *A. irregularis pentacanthus* as plain variations of *A. irregularis* (Döderlein 1917; Koehler 1924; Clark and Downey 1992). Only Ludwig (1897) used the name *A. pentacanthus* D. Ch. for *A. irregularis pentacanthus* D. Ch., thus elevating it to full specific status. While *A. bispinosus, A. jonstoni*, *A. platyacanthus* and *A. spinulosus* are considered endemic to the Mediterranean (Döderlein 1917; Tortonese 1965; Clark and Downey 1992), *A. aranciacus* and *A. irregularis* are also

reported as inhabiting the East Atlantic. Whether the variant *A. irregularis pentacanthus* is endemic to the Mediterranean, as Clark and Downey (1992) believe, or occasionally occurs in the Atlantic (Ludwig 1897; Döderlein 1917), is still uncertain.

General ecology and larval development in *Astropecten*

Representatives of *Astropecten* occur worldwide, mainly in shallow waters of tropical and temperate seas. However, several species can also live at greater depths, for instance *A. irregularis* (> 1000 m; Döderlein 1917; Tortonese 1965) and *A. hermatophilus* (up to 1500 m; Clark and Downey 1992). *Astropecten* burrow in sandy or muddy substrate where they feed on molluscs and other infaunal invertebrates (Tortonese 1965).

Most species of *Astropecten* seem to have what Mortensen (1921, 1937) considered a typical bipinnaria larva, metamorphosing in the water column and settling as young sea stars (Clark and Downey 1992). Laboratory rearing has shown that the planktonic stage usually lasts for approximately 19-25 days (Mortensen 1921; Newth 1925; Mortensen 1937; Ventura et al. 1997). Komatsu (1975), however, found a more barrel-shaped larva in *A. latespinosus* which completed metamorphosis about 5 days after fertilization. While bipinnaria larvae in *A. polyacanthus* and *A. velitaris* can also have short planktonic stages of only 3-5 days (Mortensen 1937), larval lives as long as 60 days have been observed in *A. aranciacus* (Hörstadius 1938). Although observations on larval development under laboratory conditions do not necessarily correspond to development in nature, we can assume that the planktonic stage in several species of *Astropecten* is rather long. The potential for dispersal is therefore high, and some species also have a large species range (e.g., *A. monacanthus* and *A. polyacanthus*). On the other hand, *A. polyacanthus* has a larval stage of a few days, yet has been reported from the Red Sea, the Indian Ocean and the Pacific (Döderlein 1917). This could be because the planktonic larval stage is much longer in the sea than in the laboratory, because gene flow is (or was) accomplished through continuously distributed populations, or because more than one biological species is included in *A. polyacanthus*. As food supply can strongly influence the development of larvae (Mortensen 1921), waters with a low nutrient content might prolong the duration of the larval phase and increase dispersal capacity. It is likely that larval life is not the only factor affecting species range.

Phylogenetic relationships in *Astropecten*

To date, phylogenetic relationships in *Astropecten* have not been inferred either by a cladistic approach or on the basis of molecular markers. However, in 1917 Döderlein published a monograph on the genus *Astropecten* and its evolutionary history based on morphological characters. Döderlein assigned over 100 species of *Astropecten* to different groups according to their presumed relationships. He stated that there are three major groups of species in *Astropecten* corresponding to geographic regions: 1. the East Atlantic and Mediterranean, 2. the east and the west coast of America, and 3. the Indo-West Pacific. According to Döderlein, there is generally no close relationship between the species from these different regions. Between

the East Atlantic group and the American group, he deduced deep phylogenetic separation. He thought that there was also phylogenetic separation between West African and East African *Astropecten*.

Nevertheless, morphologically similar species exist in different geographical regions and occur in similar habitats. Considering the long larval stage and potential dispersal capacity in several *Astropecten* species, the question remains open whether some of these species are closely related, or whether they have come to resemble each other due to convergent evolution. For instance, Döderlein suggested that *A. regalis* from the East Pacific and *A. marginatus* from the West Atlantic are sister species, and that a third species from Japan, *A. latespinosus*, is closely related to these two. Döderlein further proposed that in general the direction of species' spread is from east to west, and that Hawaii could serve as a stepping stone in the Pacific. The phylogenetic relationships as proposed by Döderlein are shown in Figure 1 and contain by and large the same species included in the present study. While several species with similar morphologies occur allopatrically, other presumed sister species also exist sympatrically, such as *A. articulatus* and *A. antillensis* in the West Atlantic and *A. bispinosus* and *A. platyacanthus* in the Mediterranean. Whether sympatric speciation has occurred in *Astropecten* has not yet been clarified.

Phylogenetic relationships of Mediterranean *Astropecten*

According to Döderlein (1917), Mediterranean species can be subdivided to three groups, some of which also include species from West Africa: 1. the *Irregularis* group, including *A. irregularis* and all its subspecies and variations plus *A. weberi*, a West African species; 2. the *Aranciacus* group, including *A. aranciacus*, *A. bispinosus*, *A. platyacanthu*s and *A. spinulosus*; and 3. the *Jonstoni* group, limited to *A. jonstoni* and the West African species *A. dahomensis*. He suggested that the *Irregularis* group is the most primitive, followed by the *Jonstoni* group. Döderlein also proposed that some American species, forming what he called the *Braziliensis* group, derived from species of the *Irregularis* group, and that species of the *Aranciacus* group in turn originated from the *Braziliensis* group. Within the *Aranciacus* group, Döderlein (1917) considered *A. aranciacus* as the most primitive member. Döderlein (1917) further assumed that species with fewer supero-marginal spines were generally younger than species with more supero-marginal spines, as over time this character has been reduced.

Aim of this study

The aim of this study is to resolve general taxonomic issues, identify speciation patterns and estimate species diversity within the genus *Astropecten*. We try to resolve these questions by inferring phylogenetic relationships on a global scale using molecular markers of mitochondrial DNA. Given the high diversity and extensive geographical occurrence of *Astropecten*, it is beyond the scope of this study to formally revise all described species of this genus. Nevertheless, we attempted to obtain data from as many species as possible with particular emphasis on the Mediterranean species. In particular, (1) we used a molecular phylogeny to examine the validity of the phylogenetic relationships suggested by Döderlein (1917); (2) we investigated the

possibility of morphological convergence of similar species from different geographical regions; (3) we identified potentially problematic species descriptions and/or signs of cryptic speciation in *Astropecten.*

Materials and Methods

Sampling and DNA extraction

We obtained 118 specimens belonging to 40 species of *Astropecten*. Seven additional specimens belonging to four astropectinid genera and one goniasterid genus were sampled as outgroup taxa (Table 1). Specimens were obtained by scuba diving, from trawl and gill net operations, institutional invertebrate collections and private persons from 38 locations around the world (Figure 2).

From the Mediterranean and the East Atlantic, we collected 54 specimens belonging to all six recognized species of *Astropecten* from different sampling locations as listed in Table 1. We did not distinguish between the different varieties of *A. irregularis* when sampling, as many specimens were juveniles and therefore did not exhibit the typical morphological characters necessary for this distinction.

Samples were preserved in 96% ethanol or in 80% ethanol buffered with DMSO. DNA from approximately 30 mg of arm tip tissue or tube feet was extracted using a DNeasy Tissue Kit® (QIAGEN) following the manufacturer's instructions for extraction of animal tissue for a final volume of 400 µl and then were stored at -20 °C.

DNA sequencing

We amplified fragments of three regions of mitochondrial DNA (mtDNA): approximately 576 bp of the 12S ribosomal RNA (12S), 624 bp of the 16S ribosomal RNA (16S) and 619 bp of the cytochrome c oxidase sub-region I (COI). The primers used for DNA amplification are listed in Table 2. COI was not easy to amplify in some specimens, especially those stored in museums, and thus often required various combinations of primers. DNA amplifications were performed in a 30 µl-volume reaction with 1.67 U *Taq* DNA *Polymerase*, 3 µl 10x PCR reaction buffer, 0.4 mM dNTPs, 0.2 µM of each primer, 1 mM $MgCl_2$ and 6 µl of template DNA. The PCR protocol consisted of an initial denaturation step at 96 °C for 5 s, 40 amplification cycles (95 °C for 30 s, 50 °C for 45 s and 72 °C for 1 min) and a final elongation step at 72 °C for 10 min performed in a Whatman Biometra T1 Thermocycler. The PCR products were purified with the QIAquick® PCR Purification Kit (QIAGEN) following the supplier's instructions. Forward and reverse sequencing was carried out using the primers marked with an asterisk in Table 1 and using BigDye® Terminator (PE-Applied Biosystems) chemistry. The cycle-sequencing protocol consisted of an initial step at 96 °C for 3 min and 24 sequencing cycles (96 °C for 15 s, 50 °C for 10 s and 60 °C for 3 min). Cycle

sequencing products were purified with a DyeEx™ 2.0 Spin Kit (QIAGEN) and subsequently sequenced in an ABI 3730 DNA Analyzer. Sequences were edited using the software SEQUENCHER™ 4.6 (Gene Codes Corporation) and deposited in GenBank under the accession numbers listed in Table 1.

Alignment

To align the sequences, we used the software CLUSTAL X version 2.0 (Thompson et al. 1997; Jeanmougin et al. 1998) applying the profile alignment mode. We aligned pairs with lower distances first with a gap opening penalty of 10.0 and extension penalty of 0.2. The alignment software SOAP v. 1.1b1 (Loytynoja and Milinkovitch 2001) was used to find regions that were not well supported when comparing alignments with gap opening penalties ranging from 5.0 to 15.0 in steps of two and gap extension penalties ranging from 0.1 to 10.1 in steps of two. We considered alignment regions that had less than 80% support ambiguous and removed them from the dataset using the software BIOEDIT v. 7.0.9.0 (Hall 1999). The final alignment included 366 bp of 12S, 488 bp of 16S and 546 bp of COI and can be viewed in the Appendix.

Phylogenetic analysis

We tested for phylogenetic congruence between the three mtDNA regions 12S, 16S and COI performing a partition homogeneity test as implemented in PAUP* v. 4.0b10 (Swofford 2003). This test produced a *p* value of 0.01, suggesting significantly different phylogenetic signals between the three mtDNA regions. However, phylogenetic accuracy is only negatively affected by combining data when $p < 0.001$ (Cunningham 1997). 12S, 16S and COI sequence data were, therefore, concatenated to perform phylogenetic analyses.

After removing 11 redundant haplotypes from the matrix, the program MODELTEST v. 3.7 (Posada and Crandall 1998) was used to perform hierarchical likelihood ratio tests (hLRTs) to select the best model of DNA evolution. The transversional model with a proportion of invariable sites (I) and gamma distribution (Γ) (TVM+I+Γ) was selected as the model of evolution that best fits the data (- lnL = 17244.8; α value of Γ = 0.7589; pinvar = 0.5120). Uncorrected genetic distances, as well as corrected distances using this model were calculated in PAUP*.

Global phylogeny

We performed maximum parsimony (MP) analysis in PAUP * by heuristic search for 50 random addition replicates using TBR branch swapping option and keeping 100 trees per replicate. Gaps were treated as missing data. We estimated clade support in PAUP * performing 1,000 bootstrap resampling replicates and computing decay indices (Bremer 1988).

Bayesian inference (BI) was performed in MRBAYES v. 3.1.2 (Ronquist and Huelsenbeck 2003) assuming the GTR+I+Γ model and unlinking the partitions so that parameters were estimated for each partition separately. Two runs of four chains were performed for 5,000,000 generations sampling every 500

generations and with a temperature for the heated chains of T = 0.02. After discarding the first 2,000 trees as burn-in, we used the remaining trees to estimate posterior probabilities indicating clade credibility.

Phylogeny of Mediterranean species

As the molecular phylogeny on a global scale suggested that there is possibly a monophyletic unit composed of Mediterranean and East Atlantic species, we performed a separate maximum likelihood (ML) analysis for this group. We included an additional *A. irregularis* specimen in the analysis, although only the 16S region was available in GenBank (accession code: AY652501; Kirby 2004). *Ctenopleura* sp. was used as outgroup. As the partition homogeneity test did not reveal significant inconsistencies between the three mtDNA regions for Mediterranean and East Atlantic specimens (P = 0.8), all three regions were again combined into one matrix. ML analysis was performed in PAUP * by heuristic search and 10 random addition replicates applying the TBR branch-swapping option. For the first ML search, we assumed the TVM+I+Γ model with the parameters previously estimated in MODELTEST for all specimens. Then, parameters were re-estimated on the resulting tree and used for a next ML search. This process was repeated until parameter estimates and topologies converged. Clade support was estimated with 1,000 bootstrap replicates in PAUP. We also performed a reconstruction using Bayesian inference (BI) in MRBAYES for 2,000,000 generations (burn-in 250,000) with T = 0.2 for heated chains, but otherwise applying the same settings as for the global analysis. Posterior probabilities of nodes were placed in the ML tree as additional indication of clade credibility.

Results

Specimen and sequence data

We obtained sequence data of the mitochondrial DNA regions 12S rRNA, 16S rRNA and cytochrome c oxidase subregion I (COI) of 118 specimens of *Astropecten* and seven outgroup specimens belonging to six genera. Of the roughly 40 species of *Astropecten* collected for this study, 10 could not be identified reliably neither by morphological characters nor by mtDNA sequences. Unidentifiable specimens were either juveniles lacking morphological characteristics crucial for identification, or adults that did not match any of the species descriptions in the literature. The latter was particularly the case in deep sea specimens from the South Pacific. COI amplification was not successful for 33 specimens and 16S amplification for six (see Table 1 for missing GenBank accession numbers). Of the 1390 base pairs of the concatenated sequence from the three mtDNA regions, 792 were invariant and 508 were potentially phylogenetically informative.

Global phylogeny

The maximum parsimony (MP) analysis on a global scale for all mtDNA regions combined resulted in 800 most parsimonious trees (length = 3564), from which we constructed a strict consensus tree. This total

evidence tree with bootstrap support values (BSP) and decay indices (DI) is shown in Figure 3a. Ingroup taxa were monophyletic (BSP = 93; DI = 18) and clustered in three main clades, each clade consisting of taxa from the same geographic region. The three clades corresponded to the following regions: 1. Indian Ocean and Pacific, excluding the West American coast (clade A); 2. East and West coast of America (clade B); and 3. East Atlantic and Mediterranean (clade C).

Bayesian inference (BI) resulted in 10,000 trees of which the first 2,000 were discarded as burn-in. The remaining trees were used to construct a consensus tree following the 50 % majority rule. BI exhibited a tree topology similar to the maximum parsimony (MP) tree and was not in conflict with respect to any of the major clades (Figure 3b). Posterior probability (PP) supporting the monophyly of the ingroup equalled 100%, and the three main geographic regions within the ingroup were again distinct; however, this time *A. aranciacus* specimens appeared as a separate, fourth clade in the phylogeny and did not group with the remaining Mediterranean and East Atlantic species.

Clade A: Indo-Pacific

Although the Indo-Pacific clade was not supported by bootstrap values (BSP <50) or decay index (DI = 1), nodal support was provided by posterior probability (PP = 94) (Figure 3). Within this clade, three main groups formed: the first clade contained all *A. polyacanthus* specimens, except for one from Fiji (clade D; BSP = 97; DI = 18; PP = 100); the second clade comprised a mix of species from various geographic regions ranging from the North Arabian Sea to the South Pacific (clade E; BSP < 50; DI = 11; PP = 89); and the third clade included specimens exclusively from the South Pacific Islands (clade F; BSP = 52; DI = 5, PP = 93). Within clade D, *A. scoparius* from Japan and an unidentified specimen (*A.* sp. 1) from Fiji were sister to the clade comprising mainly *A. polyacanthus* (clade G). *A. polyacanthus*, however, was paraphyletic, as two other species, *A. triseriatus* from Hawaii and *A. vappa* from Brunei, grouped within this clade. *A. triseriatus* grouped with *A. polyacanthus* from Hawaii, Japan and Dubai (BSP = 87; DI = 8; PP = 100) and *A. vappa* with *A. polyacanthus* from New Zealand (BSP = 92; DI = 7; PP = 99). Between specimens in the *A. polyacanthus* clade, uncorrected genetic distances ranged from 0.003 to 0.107 substitutions per site. Corrected genetic distances using the parameters estimated in MODELTEST were between 0.002 and 0.143 substitutions per site.

Clade E included various species from different locations and also one *A. polyacanthus* specimen from Fiji. Specimens of *A. indicus* from Brunei, Pakistan and Thailand formed a monophyletic clade (BSP = 78; DI = 2; PP = 100). Between specimens of *A. indicus* corrected and uncorrected genetic distances ranged from 0.001 to 0.054 and from 0.001 to 0.066 substitutions per site, respectively.

Specimens grouping into clade F were exclusively from the South Pacific and were collected at a greater depth. Specimen Asp2-Fij1 was collected at ca. 250 m, and the rest at over 320 m. Two other specimens from Fiji, *A. polyacanthus* (Apoly-Fij1) and an unidentified specimen (Asp1-Fij1), collected at <180m, did not appear in this clade, but grouped with clade E.

Clade B: East and West coast of America

All specimens from the East and the West coast of America grouped into one clade confirmed by high node support (BSP = 98; DI = 23; PP = 100). Within this clade, specimens from the Pacific coast clearly grouped together (clade H; BSP = 99; DI = 13; PP = 100). *A. verrilli* was included in this clade, but it was not monophyletic. Between *A. armatus* and *A. siderealis* corrected and uncorrected genetic distances equalled 0.003 substitutions per site, indicating that these specimens very likely belong to the same species. Species from the Atlantic coast were paraphyletic; they grouped into three different clades according to the MP tree (clades I, J and K). Clade I included *A. alligator*, *A. americanus*, *A. comptus* and *A. nitidus* and was only well supported by PP (BSP = 51; DI = 5; PP = 92). Specimens of *A. alligator* and *A. americanus* probably belong to the same species, as genetic distance between the two equalled 0.001 substitutions per site (corrected and uncorrected). *A. antillensis*, *A. duplicatus* and three specimens of *A. articulatus* grouped into clade J (BSP = 98; DI = 7; PP = 100), but *A. articulatus* was not monophyletic, because two additional specimens formed a clade sister to J. The third group, clade K, was only supported by DI (11) and consisted of *A. marginatus* and *A. cingulatus*.

Clade C: Mediterranean and East Atlantic

Maximum parsimony and Bayesian inference.

MP analysis and BI were consistent in considering *A. aranciacus* as a reciprocally monophyletic clade with respect to the other Mediterranean and East Atlantic species. However, the two methods differed with regard to the relation of the *A. aranciacus* clade to the rest of the Mediterranean and East Atlantic species. BI indicated that *A. aranciacus* is a completely independent clade (clade C2), whereas MP grouped *A. aranciacus* with the other species from this region in clade C (Figures 3 and 4), albeit only with weak support (BSP < 50; DI = 9). Each Mediterranean and East Atlantic species was monophyletic except for *A. irregularis*. This species was paraphyletic, as Portuguese specimens (clade L; Far1-3) appeared separated from the rest of the species (clade M). Non-Portuguese *A. irregularis* were more closely related to *A. spinulosus* (clade N; BSP = 72; DI = 11; PP = 97) than to the Portuguese *A. irregularis*. While *A. bispinosus* and *A. platyacanthus* were sister species (BSP = 100; DI = 18; PP = 100) and related to the *irregularis-spinulosus* group (BSP = 58; DI = 9; PP = 76), sequences of *A. jonstoni* and *A. africanus* were more distantly related to the rest.

Maximum likelihood analysis of Mediterranean and East Atlantic species:

Maximum likelihood (ML) analysis for only Mediterranean and East Atlantic specimens yielded one tree with a likelihood -lnL of 5480.33 (molecular clock not enforced) as shown in Figure 5. The tree exhibited a similar topology as the MP and BI consensus trees for the same specimens. *A. irregularis* from Portugal (Far1,2,3; Clade L) were again clearly separated from the rest of the *A. irregularis* sequences (Clade M), which grouped with *A. spinulosus* (clade N; BSP = 91; PP = 100). However, within the non-Portuguese *A. irregularis* there was clear evidence for two separate clades, clade O (BSP = 78; PP = 100) and clade P (BSP

= 94; PP = 100). Clade P consisted of one *A. irregularis* specimen from the Irish Sea (Nor1) and one specimen from the West Mediterranean (Sar3).

The molecular phylogeny thus suggests that *A. irregularis* consists of at least three subclades (clades L, O and P) and is paraphyletic, because it includes *A. spinulosus*. Uncorrected genetic distances within subclades were below 0.015 (corr. < 0.016) substitutions per site, whereas distances within each of clades L and P were below 0.003 substitutions per site (uncorr. and corr.; see Table 3). Uncorrected and corrected genetic distances within the other Mediterranean and East Atlantic species were below 0.090. Between Portuguese and non-Portuguese *A. irregularis* genetic distances ranged from 0.053 to 0.084 (uncorr.) and from 0.062 to 0.111 (corr.) substitutions per site. Genetic distances between clade O and clade P of non-Portuguese *A. irregularis* were between 0.016 and 0.038 (uncorr.) and between 0.017 and 0.042 (corr.) substitutions per site.

In summary, our results showed evidence of a phylogenetic separation between species from the three geographic regions: 1. Indo-Pacific; 2. Neotropics; and 3. Mediterranean and East Atlantic. The total evidence molecular phylogeny did not support any close relationship among species of these three different regions. Only the Mediterranean and East Atlantic *A. aranciacus* maintained an unresolved position in this respect, as it did not confidently cluster together with other species of the Atlanto-Mediterranean region or with species of any other region. Furthermore, several species were not monophyletic according to molecular data, including the Mediterranean and East Atlantic species *A. irregularis*.

Discussion

Independent of phylogenetic algorithm, molecular phylogenies displayed a phylogenetic gap between *Astropecten* species of three geographic regions: 1. Indo-Pacific; 2. Neotropics; and 3. East Atlantic and Mediterranean. The tree topologies resulting from maximum parsimony (MP) analysis and Bayesian inference (BI) were mostly in agreement but were not conclusive in respect to the position of *A. aranciacus*. As is often the case with mtDNA sequence data, deeper nodes received lower bootstrap support, and relationships at this level were not as well resolved. Nevertheless, posterior probabilities were high, placing the three geographic regions into different phylogenetic units. To obtain better resolution and higher statistical support at deeper nodes, sequencing of a slower evolving gene, such as the nuclear 18S, could be appropriate. However, as we were working with some very old specimens with low DNA content and fragmented DNA, this would have required additional resources.

While the mitochondrial 12S and 16S regions amplified successfully in almost all cases, we encountered some difficulties when amplifying the COI region, particularly in older museum specimens. The quality of the DNA and the variability in this region required several different primer combinations and did not always amplify the complete fragment. Although we were not able to obtain the COI sequence for several specimens

and, in a few cases, the 16S sequence, we included all taxa in the analysis. Adding taxa with missing data presumably affects phylogenetic resolution more positively than excluding them (Wiens 2006). While in model-based analysis, such as BI, even taxa with highly incomplete sequences (> 75% missing data) improve phylogenetic accuracy, in MP more than 25% of the data per sequence are required to rescue the analysis from long branch attraction (Wiens 2006). As the 12S region amounts to more than 25% of the included base pairs, and as Bayesian inference resulted in a very similar tree topology, we think that there were no deleterious effects from missing data in any of the analyses.

Molecular phylogeny vs. the phylogeny of Döderlein, 1917

As Döderlein (1917) had suggested on the basis of morphology, there is a phylogenetic gap between species that inhabit three large geographic regions. However, Döderlein also assumed that some extant species can be seen as connected between these regions, as for example *A. marginatus* from the Atlantic and *A. regalis* from the Pacific coast of America. Döderlein proposed that these two species are sisters and are closely related to *A. latespinosus* from Japan, thus acting as a "phylogenetic connection" between these regions. The molecular phylogeny does not support this hypothesis as each of these three species clearly groups with other species of the respective geographical region. Thus, molecular data indicate that the similarities of these species in morphology and habitat preference are likely to have evolved independently.

While the general grouping of the three large geographic regions as Döderlein suggested is in agreement with molecular data, there are several differences on a more local scale:

Indo-Pacific region:

Within the Indo-Pacific clade the only agreement of the molecular phylogeny with Döderlein's grouping is that most *A. polyacanthus* specimens cluster into the same clade, and that *A. monacanthus* and *A. granulatus* are closely related. However, *A. indicus* is not in the same group with the latter two species, as Döderlein had suggested (see Figure 1). Moreover, although *A. vappa* and *A. triseriatus* both cluster within the *Polyacanthus* clade, they are phylogenetically not as closely related as Döderlein proposed. Also, according to Döderlein, *A. javanicus* would be in the same group as *A. polyacanthus*, but molecular data place this species with *A. zebra,* which Döderlein considered as belonging to the *Velitaris* group.

American region:

MtDNA shows evidence of a phylogenetic separation between East Pacific and West Atlantic species. This generally agrees with Döderlein's assumptions, except that molecular data do not show evidence of a close relationship between the West Atlantic *A. marginatus* and the East Pacific *A. regalis* as Döderlein had proposed. Among the West Atlantic species there are similarities between Döderlein's *Articulatus* group and the clade comprised of *A. antillensis*, *A. articulatus* and *A. duplicatus* in the molecular phylogeny. However, Döderlein's inclusion of *A. cingulatus* in this group disagrees with the molecular data, which either grouped *A. cingulatus* with *A. marginatus,* or placed it in a separate clade. Similarly, Döderlein assigned *A.*

americanus to the *Articulatus* group, but molecular data place this species in a different clade together with *A. alligator*, *A. nitidus* and *A. comptus*. Döderlein did not include the latter three species in his monograph.

Döderlein also suggested that Hawaii is a stepping stone for species spreading from East to West, which he believed to be the general direction of species colonization around the globe. The molecular data show no evidence of this connection, as all Hawaiian specimens clearly group within the Indo-Pacific clade. Similar results have been found in other marine taxa such as sea urchins and fish (Colborn et al. 2001; Lessios et al., 2001). It is more probable that the *Astropecten* phylogeny was mainly shaped by geological events, such as the separation of the American and the African continents, the closure of the Tethys, the appearance of the Benguela upwelling, and the rise of the Isthmus of Panama.

Mediterranean and East Atlantic:

Our molecular and Döderlein's approaches agree that *A. bispinosus* and *A. platyacanthus* are sister species and place *A. irregularis* in a separate clade. On the other hand, the molecular phylogeny includes *A. spinulosus* in the *Irregularis* clade, whereas Döderlein considered *A. spinulosus* to be more closely related to *A. bispinosus* and *A. platyacanthus* within the *Aranciacus* group. According to Döderlein, *A. jonstoni* forms a group separate from the other Mediterranean species. Our results support this assumption with a relatively long branch leading to this species in the maximum likelihood (ML) tree, and a topology indicating that this species separated from the other Mediterranean species before subsequent radiation. The main difference to Döderlein's phylogeny is that mtDNA sequence data place *A. aranciacus* at the base or even outside of the Mediterranean and East Atlantic group and not in the same group as *A. bispinosus* and *A. platyacanthus*. According to the molecular phylogeny, *A. aranciacus* was the first species to split off from the other Mediterranean species and would deserve a group of its own in Döderlein's system. Furthermore, mtDNA did not place *A. africanus* within the *Irregularis* group as Döderlein proposed, but rather suggests that this species diverged earlier from the other species of the *Irregularis* group. Based on molecular data it is therefore not appropriate to consider *A. africanus* a subspecies of *A. irregularis*.

Contrary to Döderlein's view, the molecular phylogeny neither supports a phylogenetic relationship between the *Irregularis* group (clades L and M) and the *Braziliensis* group (clade H) nor suggests any relationship between the *Braziliensis* group and the *Aranciacus* group (clade C).

In summary, although the comparison between the molecular phylogeny and Döderlein's morphological relationships reveals a great deal in common in the large scale relationships of geographic groups, many discrepancies emerge on a local level. Assessing morphological characters and then using a cladistic approach could perhaps result in a more adequate comparison of molecular and morphological phylogenies. However, morphological diversity is very high in *Astropecten* and characters are often continuous rather than discrete. Also, many characters are only expressed in fully grown adults and not in juvenile specimens. For these reasons, building a matrix of morphological characters is not easy and would probably lead to many

ambiguities. Considering the morphological complexity, we believe that until meaningful characters and character categories have been determined, molecular data are a more reliable approach to resolve phylogenetic relationships in *Astropecten* than morphological characters.

Taxonomic issues

Although it is beyond the scope of this study to revise the genus *Astropecten*, our results suggest several reassessments of the current taxonomy:

Astropecten polyacanthus Müller and Troschel

Given the genetic variation within this clade, it would be more appropriate to speak of a *Polyacanthus* species complex. Although the genetic distances of > 10% substitutions per site, particularly between geographically distant specimens, suggest separate species, the morphology is very similar, and current descriptions of *Astropecten* do not permit delineation of species. By comparison, the sister taxa *A. articulatus* and *A. antillensis,* which are morphologically distinct, show a genetic distance of around 1.8% substitutions per site for the studied mtDNA regions. The specimen from Dubai, which is morphologically most similar to *A. polyacanthus,* actually lacks the typical character of missing spines on the 2. (- 4.) supero-marginal plate. Nevertheless, according to molecular data, it qualifies for inclusion in the *Polyacanthus* species complex. On the other hand, a specimen from Fiji expressing this character, did not group within this complex in the molecular phylogeny. Our results show that current morphological descriptions are not sufficient to distinguish between the species, and that cryptic speciation is most likely present in this group.

Astropecten indicus Döderlein

This species appears monophyletic in the molecular phylogeny; however, specimens from Brunei, Pakistan and Thailand are genetically clearly distinct from each other with distances over 5% substitutions per site. Therefore, our data suggest that specimens of this clade are again part of a species-complex and morphological descriptions need to be refined, because cryptic speciation might have also taken place in *A. indicus*.

Astropecten verrilli de Loriol

Although *A. verrilli* is mainly described from West coast of Central America and *A. californicus* (Fisher 1906) from California, Döderlein (1917) synonymised the two as there are no apparent morphological differences between these two species. Some of our specimens from the Pacific coast of Panama also meet the descriptions of *A. verrilli*, but in the molecular phylogeny they appear in different clades than the Californian specimens. Therefore, our data suggest that *A. verrilli* and *A. californicus* are not synonyms and that *A. californicus* should be used for the specimens that were collected in San Diego, California (Cal2, Cal3). This implies that the specimen from Monterrey (Cal1) belongs to yet another species, but its small size prevented us from using morphological criteria for assigning it to a known species.

In general, there is a great deal of confusion regarding the taxonomy of *Astropecten* from the West American coast (Fisher 1906; Verrill 1914; Döderlein 1917; Ziesenhenne 1939). Several characters that have been used to describe species in this area (such as the presence of supero-marginal spines and number of paxillae per plate) are not always expressed in juveniles. Descriptions are often based on preserved specimens alone - thus not including any indication of color when alive. They are generally not sufficient to identify specimens reliably.

Astropecten articulatus Say

A. articulatus is not monophyletic in the molecular phylogeny. Specimens obtained from the University of South Carolina have only very small spines on the supero-marginal plates and can therefore be assigned to the variation *A. articulatus* var. valenciennii Döderlein. As these specimens are not sister to *A. articulatus* from Panama, and as the genetic distance between the two is above 3% substitutions per site (corr. and uncorr.), it would be appropriate to consider them as a separate species rather than a variation.

Astropecten bispinosus Otto and *Astropecten platyacanthus* Philippi

The molecular phylogeny and the genetic distances between specimens clearly indicate that *A. bispinosus* and *A. platyacanthus* are not just variations, as Ludwig (1897) and Koehler (1924) thought, but should be considered as separate species. Although these two species occur sympatrically, they are morphologically clearly distinguishable. However, some intermediate forms have been found (G. Ribi et al. unpublished), suggesting that occasional hybridization between the two sister species is a possibility.

Astropecten irregularis Pennant and *Astropecten pentacanthus* Delle Chiaje

Among *A. irregularis* specimens, genetic variability is high. Our data suggest that at least three species can be distinguished. Specimens from Portugal (Far1, 2 and 3), collected at a much greater depth (100 - 540 m), are genetically clearly distinct from other *A. irregularis*. Although the Portuguese specimens meet the description of *A. irregularis pentacanthus* Delle Chiaje, this description may apply to more than one species. *A. irregularis* has been recorded from depths of over 900 m, but it is not clear whether specimens from the deep belong to the same species as the ones collected in Portugal. Genetic data of *Astropecten* from deeper waters of the Mediterranean and North-East Atlantic could provide valuable information on diversity within *Astropecten*. *A. ibericus* is another species recorded from the Iberian Peninsula from depths reaching 130 m, and is similar in morphology to the specimens from Portugal. However, as only small specimens of *A. ibericus* have been collected so far, Clark and Downey (1992) considered this species doubtful. Genetic data of *A. ibericus* could clarify the validity of this species and would resolve the relation to the Portuguese specimens used in this study. Within non-Portuguese *A. irregularis*, there is evidence for two clades. Although these clades are sister to each other, the genetic distance of 2% between them is large enough to raise the possibility of separate specific status. One of these clades includes one specimen from the North Sea (Nor1) and one from Sardinia (Sar3). *A. irregularis* from the North Sea typically carries a larger conical

spine on each of the upper marginal plates and has been referred to as *A. irregularis irregularis* by Döderlein (1917) and by Clark and Downey (1992). The Mediterranean form has been established in the literature as *A. irregularis pentacanthus* and lacks spines on the supero-marginals (Döderlein 1917; Tortonese 1965; Clark and Downey 1992). Contrary to the opinion of some taxonomists that the typical North Sea and the typical Mediterranean form are merely local varieties (Döderlein 1917; Clark and Downey 1992), genetic differences suggest that these "varieties" should be treated as two separate species. Adult specimens carrying spines on the upper marginal plates should therefore be named *A. irregularis* and adults lacking spines on the upper marginals *A. pentacanthus*. *A. pentacanthus* has already been used as a species name by Müller and Tröschel (1842) and Ludwig (1897) and was first described as *Asterias pentacanthus* by Delle Chiaje (1825). Our results demonstrate that *A. pentacanthus* and *A. irregularis* occur sympatrically in the Mediterranean and that *A. pentacanthus* also occurs in the Atlantic as indicated by the specimen from Madeira (Mad1). To assess the distribution of *A. irregularis* in the Mediterranean and *A. pentacanthus* in the Atlantic, further data are needed.

Conclusions

Phylogenetic inferences based on mtDNA indicate that morphological convergence has taken place in *Astropecten* resulting in allopatric non-sister taxa with similar morphologies and habitat preferences. Although morphology suggests several close relationships between species in geographically distant areas, molecular data show evidence of a clear phylogenetic separation of these regions. The comparison to Döderlein's morphological phylogeny reveals many discrepancies, particularly on a local scale, indicating that informative morphological characters are not easily identified and categorized in *Astropecten*.

Döderlein assumed, based on morphological characters, that *A. marginatus* and *A. regalis* on either side of the Isthmus of Panama were closely related. The molecular phylogeny presents a different reconstruction of evolutionary history. *A. marginalis* is, in fact, distantly related to *A. regalis*. Instead, a transisthmian pair is formed by clades I (West Atlantic) and H (East Pacific) (Figure 3). This suggests that the rise of the Isthmus of Panama, which was completed approximately 3 mya, separated populations of the common ancestor of these two clades, which then split into *A. duplicatus, A. antillensis,* and *A. articulatus* in the Atlantic and into *A. verilli*, *A. erinaceus, A. regalis, A. oerstedi*, *A. siderealis,* and *A. armatus* in the eastern Pacific. For the entire concatenated sequence of 12S, 16S and COI, the uncorrected divergence between clades I and H in *Astropecten* is 6.3% and the corrected divergence is 7.4%. For separate genes, the uncorrected genetic distance was 4.8% in 12S, 5.2% in 16S and 12.3% in COI. Other echinoderms likely to have been split by the rise of the Isthmus, show divergences of 5.6-6.1% in 12S, 6.5-12.8% in 16S, and 8.7-13.5% in COI (Lessios 2008). It is, therefore, quite possible that clades I and H are geminate, having been split not much more than 3 mya. If so, the generation of so many species on either side of the Isthmus within

this relatively short period of time indicates a remarkable rate of speciation, unparalleled by any other echinoderm with the same presumed phylogenetic history relative to the Isthmus. Among other echinoderms with known transisthmian clades, only the echinoid *Lytechinus* comes close, with three post-isthmian species in the eastern Pacific and three (or possibly four) in the Atlantic (Zigler and Lessios 2004).

The molecular data indicate that it is appropriate to consider several widely distributed taxa, such as *A. polyacanthus* and *A. indicus*, as species complexes. Many formerly presumed within species morphological variants exhibit genetic distances which are large enough to raise them to the species level. For instance, our data suggest that *A. irregularis* var. *pentacanthus* can be considered as *A. pentacanthus*. In several other cases, such as *A. polyacanthus* and *A. indicus*, the possibility of cryptic speciation within each of these species cannot be ruled out and requires additional investigations.

In a few cases genetic data support the view that some described species should be synonymized. However, many new species remain to be described, such as several deep-sea populations of the South Pacific. This suggests that even though the genus *Astropecten* is known for its species-richness, the diversity in this genus might yet be underestimated. In *A. indicus* and *A. polyacanthus* sibling species occur allopatrically and most likely have speciated by vicariance. Based on current distributions of sister species, sympatric speciation in *Astropecten* cannot be ruled out. The possibility of sympatric speciation has also been considered for sea urchins, such as *Diadema* (Lessios et al. 2001) and *Lytechinus* (Zigler and Lessios 2004), and although it has been rejected in *Diadema,* it is not clear whether all *Lytechinus* species are the result of allopatric speciation. In the Indo-Pacific region molecular data reveal a few close relationships between supposedly conspecific populations of *Astropecten* from geographically distant locations such as between *A. polyacanthus* specimens from Hawaii and from Dubai. Waters et al. (2004) suggested that large distance distribution throughout the Indo-Pacific even occurs in some asterinids lacking planktonic larval stages. To what extent ocean currents or other mechanisms have lead to this pattern in South Pacific *Astropecten* remains to be determined.

Many difficulties were encountered in this study in identifying species morphologically based on species descriptions, and it became evident that a taxonomic revision of this genus is urgent. The present study has shown that molecular markers provide a valuable complement to our predecessors' morphological work and help to clarify unresolved issues of evolutionary history and systematics.

Acknowledgements

For providing samples and hosting museum visits we thank the following people and institutions: The late C. Ahearn and the Smithsonian National Museum of Natural History, N. Améziane and the Muséum

National d'Histoire Naturelle BIMM in Paris, the Zoological Museum Amsterdam, G. Borrero and INVEMAR in Colombia, M. Brogger, A. Cabrinovich and the London Natural History Museum, T. Cerveira Borges and the BIOPESCAS team (CCMAR, University of Algarve), the California Academy of Science Collections, R. Collin, J. Collins, T. Danckwart, S. Diliberto, A. Gebruk and the Shirshov Institute, L. Granek, S. Godwin and the Bishop Museum, G. Hendler and the Natural History Museum of Los Angeles County, J. Hoover, M. Huber, M. Komatsu, D. Lane, L. Lovell and the Scripps Institution of Oceanography, C. Lüter and the Zoological Museum Berlin, H. Maag, M. Matsubara, P. Morais, C. Nielsen and the Zoological Museum of the University of Copenhagen, G. Pauley and the Florida Museum of Natural History, the Phuket Marine Biological Center, C. Ruckstuhl, H. Sekiguchi, M. Sewell, F. Solis-Marin and the Colleccion National de Equinodermos de la Universidad Nacional Autonoma de Mexico, the collections of the University of South Carolina, Q. Tahera and the Marine reference collection & Resource centre in Karachi, S. Tanner, S. Tassis, M. von Arx, K. Wakabayashi and P. Wirtz. T. Bucher, A. and L. Calderon, S. Sbilordo and S. Tanner we thank for technical and lab assistance. We are grateful for the valuable advice on data analysis we received from M. Bernasconi, A. Faucci, A. Hiller-Galvis, R. Sponer and T. Wilson. We would also like to thank D. Blake, A. Gale and C. Mah for their help with species identification and general advice. We are also greatly indebted to G. Ribi who helped outlining the project and provided assistance in many ways. This project was partially funded by the Smithsonian Tropical Research Institute and the Swiss National Science Foundation.

References

Agassiz A (1877) North American starfishes. Mem Mus Comp Zool **5**: 1-136.

Bremer K (1988) The Limits of Amino-Acid Sequence Data in Angiosperm Phylogenetic Reconstruction. Evolution **42**: 795-803.

Clark AM, Downey ME (1992) Starfishes of the Atlantic. Chapman & Hall, London, pp 794.

Colborn, J., Crabtree, R.E., Shaklee, J.B., Pfeiler, E., Bowen, B.W. (2001) The evolutionary enigma of bonefishes (*Albula* spp.): cryptic species and ancient separation in a globally distributed shorefish. *Evolution*, 55, 807–820.

Cunningham CW (1997) Can three incongruence tests predict when data should be combined? Mol Biol Evol **14**: 733-740.

Delle Chiaje S (1825) Memorie sulla storia e notomia delli animali senza vertebre del Regno di Napoli, Napoli, **2**: 185-444.

Döderlein L (1917) Die Asteriden der Siboga-Expedition. I. Die Gattung *Astropecten* und ihre Stammesgeschichte. In: Brill EJ (ed) Siboga-Expeditie. Uitkomsten op zoölogisch, botanisch, ozeanographisch en geologisch gebied verzameld in Nederlandsch Oost-Indie 1899-1900 aan boord H.M. "Siboga". 46 (a), Leiden, pp 191.

Fisher WK (1906) New Starfishes from the Pacific Coast of North America. Proc Wash Acad Sci **8**: 11-139.

Fisher WK (1911) Asteroidea of the North Pacific and adjacent waters. Part 1. Phanerozonia and Spinulosa. Bull U.S. Nat Mus **76**, pp 420.

Fisher WK (1913) Four new genera and fifty-eight new species of starfishes from the Philippine Islands, Celebes, and the Moluccas. Proc U.S. Nat Mus **43**: 599-648.

Gray JE (1841) A synopsis of the genera and species of the class Hypostoma (Asterias, Linnaeus). Ann Mag Nat Hist **1**: 175-184.

Hall TA (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucl Acids Symp Ser **41**: 95-98.

Hörstadius S (1938) Über die Entwicklung von *Astropecten aranciacus* L. Pubbl Stn Zool Napoli **17**: 221-312.

Jeanmougin F, Thompson JD, Gouy M, Higgins DG, Gibson TJ (1998) Multiple sequence alignment with Clustal x. Trends Biochem Sci **23**: 403-405.

Koehler R (1909) Echinodermes provenant des campagnes du yacht Princesse-Alice. Resultats des campagnes scientifiques (Monaco) **34**: 1-317.

Koehler R (1910) Shallow-Water Asteroidea. Echinoderma of the Indian Museum, Calcutta, Indian Museum **191**, pp 206.

Koehler R (1924) Les Echinodermes des Mers D'Europe. Librarie Octave Doin, Paris pp 210.

Komatsu M (1975) Development of Sea-Star, *Astropecten-Latespinosus* Meissner. Biol Bull **148**: 49-59.

Lessios HA, Kessing BD, Pearse JS (2001) Population structure and speciation in tropical seas: Global phylogeography of the sea urchin Diadema. Evolution **55**: 955-975.

Loytynoja A, Milinkovitch MC (2001) SOAP, cleaning multiple alignments from unstable blocks. Bioinformatics **17**: 573-574.

Ludwig H (1897) Die Seesterne des Mittelmeeres. Fauna und Flora des Golfes von Neapel und der angrenzenden Meeres-Abschnitte. Zool Stn Neapel **24**, pp 491.

Mortensen T (1921) Studies of the development and larval forms of echinoderms. Copenhagen, pp 253.

Mortensen T (1937) Contributions to the study of the development and larval forms of echinoderms III. Mémoires de l'Académie Royale des Sciences et des Lettres de Danemark, Copenhagen, pp 65.

Müller J, Troschel FH (1842) System der Asteriden. Bieweg & Sohn, Braunschweig, pp 132.

Newth HG (1925) The early development of *Astropecten irregularis*, with remarks on duplicity in Echinoderm larvae. Quart Journ Micr Sci **69**: 519-542.

Palumbi SR (1992) Marine Speciation on a Small Planet. Trends Ecol Evol **7**: 114-118.

Palumbi SR (1994) Genetic-Divergence, Reproductive Isolation, and Marine Speciation. Ann Rev Ecol Syst **25**: 547-572.

Paulay G, Meyer C (2006) Dispersal and divergence across the greatest ocean region: Do larvae matter? Integr Comp Biol **46**: 269-281. doi: 10.1093/icb/icj027.

Perrier E (1875/6) Révision de la collection de Stellerides du Muséum d'Histoire Naturelle de Paris. Arch Zool Exp Gen **4**: 265-450.

Posada D, Crandall KA (1998) MODELTEST: testing the model of DNA substitution. Bioinformatics **14**: 817-818.

Riginos C, Victor BC (2001) Larval spatial distributions and other early life-history characteristics predict genetic differentiation in eastern Pacific blennioid fishes. Proc Roy Soc B **268**: 1931-1936.

Ronquist F, Huelsenbeck JP (2003) MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics **19**: 1572-1574.

Say T (1825) On the species of the Linnean genus Asterias inhabiting the coast of the United States. J Acad Nat Sci Phil **5**: 151-154.

Sladen WP (1889) Report on the Asteroidea. Report on the scientific results of the voyage of the H. M. S. Challenger during the years 1873-1876. Zoology **30**: 1-893.

Sponer R, Roy MS (2002) Phylogeographic analysis of the brooding brittle star *Amphipholis squamata* (Echinodermata) along the coast of New Zealand reveals high cryptic genetic variation and cryptic dispersal potential. Evolution **56**: 1954-1967.

Swofford DL (2003) PAUP*. Phylogenetic Analysis Using Parsimony (*and other Methods). Version 4, Sinauer Associates, Sunderland, Massachusetts

Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG (1997) The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res **25**: 4876-4882.

Tortonese E (1965) Echinodermata Fauna d'Italia VI, Calderini, Bologna, pp 422

Ventura CRR, Falcao APC, Santos JS, Fiori CS (1997) Reproductive cycle and feeding periodicity in the starfish *Astropecten brasiliensis* in the Cabo Frio upwelling ecosystem (Brazil). Invert Reprod Dev **31**: 135-141.

Verrill AE (1914) Monograph of the shallow-water Starfishes of the North Pacific Coast from the Arctic Ocean to California. Harriman Alaska Series **14**: 1-420.

Verrill AE (1915) Report on the Starfishes of the West Indies, Florida, and Brazil. Bull Lab Nat Hist **7**: 2-232.

Waters JM, O'Loughlin PM, Roy MS (2004) Molecular systematics of some Indo-Pacific asterinids (Echinodermata, Asteroidea): does taxonomy reflect phylogeny? Mol Phylogenet Evol **30**: 872-878. doi: 10.1016/j.ympev.2003.08.019.

Wiens JJ (2006) Missing data and the design of phylogenetic analyses. J Biomed Inform **39**: 34-42.

Wilson NG, Hunter RL, Lockhart SJ, Halanych KM (2007) Multiple lineages and absence of panmixia in the "circumpolar" crinoid *Promachocrinus kerguelensis* from the Atlantic sector of Antarctica. Mar Biol **152**: 895-904. doi: 10.1007/s00227-007-0742-9.

Ziesenhenne FC (1939) The Templeton Crocker Expedition. X. Echinoderms from the West Coast of Lower California, the Gulf of California and Clarion Island. Zoologica **22**: 209-239.

Zigler KS, Lessios HA (2004) Speciation on the coasts of the new world: Phylogeography and the evolution of bindin in the sea urchin genus Lytechinus. Evolution **58**: 1225-1241.

Table 1: Species, identified code, localities and sampling date, voucher and GenBank accession codes of *Astropecten* and outgroup specimens.

Species	ID	Location/Date	Voucher	Genbank accession codes		
				12S rRNA	16S rRNA	COI
A. africanus	Aafri - STo1	EA - São Tomé/ Feb 2006		FJ171765	FJ177591	FJ195695
A. alligator	Aalli - Col1	WA - Santa Maria, Colombia; 300 m/ 2001	INV. EQU01809	FJ171787	FJ177548	
A. americanus	Aamer - Flo1	WA - Tampa Bay, Florida, USA; 271 m/ Mar 2003	UF 3471	FJ171786	FJ177547	
A. antillensis	Aanti - Pan1	WA - San Blas, Panama/ Feb 2002		FJ171794	FJ177541	FJ195654
	Aanti - Col1	WA - Arboletes, Colombia; 21 m/ 2001	INV. EQU01719	FJ171793	FJ177541	FJ195714
A. aranciacus	Aaran - Sar1	WM - Costa Colostrai, Sardinia; 17 m/ Aug 2002		FJ171773	FJ177596	FJ195679
	Aaran - Sar2	WM - Costa Colostrai, Sardinia; 17 m/ Aug 2002		FJ171774	FJ177598	FJ195681
	Aaran - Sar3	WM - Costa Colostrai, Sardinia; 17 m/ Aug 2002		FJ171775	FJ177599	FJ195683
	Aaran - Kav1	EM - Kavala, Greece; 8-18 m/ Mar 2006		FJ171770	FJ177592	FJ195669
	Aaran - Kav2	EM - Kavala, Greece; 8-18 m/ Mar 2006		FJ171766	FJ177592	FJ195668
	Aaran - Kav3	EM - Kavala, Greece; 8-18 m/ Mar 2006		FJ171767	FJ177593	FJ195669
	Aaran - Far1	EA - Lagos, Faro, Portugal; 34-42 m/ May 2005		FJ171771	FJ177597	FJ195670
	Aaran - Far2	EA - Lagos, Faro, Portugal; 34-42 m/ May 2005		FJ171772	FJ177596	FJ195671
	Aaran - Far3	EA - Faro, Faro Portugal; 30 m/ Jun 2005		FJ171770	FJ177596	FJ195669
	Aaran - Mad1	EA - Quinta do Lorde, Madeira; 15-25 m/ Oct 2005		FJ171768	FJ177594	
	Aaran - Mad2	EA - Quinta do Lorde, Madeira; 15-25 m/ Oct 2005		FJ171769	FJ177595	
	Aaran - Mad3	EA - Quinta do Lorde, Madeira; 15-25 m/ Oct 2005		FJ171777	FJ177601	FJ195729
	Aaran - Cre1	EM - Croatia, Cres; 5 m/ Oct 2002		FJ171776	FJ177600	FJ195687
	Aaran - Cre2	EM - Croatia, Cres; 5 m/ Oct 2002		FJ171739	FJ177599	FJ195674
	Aaran - Cre3	EM - Croatia, Cres; 5 m/ Oct 2002		FJ171770	FJ177596	FJ195675
A. armatus	Aarma - Mex1	EP - Puerto Peñasco, Northern Sea of Cortez, Mexico; 69 m/ Mar 1985	UNAM 4190	FJ171785	FJ177563	
A. articulatus	Aarti - Pan1	WA - Bocas del Toro, Panama/ 1996		FJ171791	FJ177543	FJ195658
	Aarti - Pan2	WA - Bocas del Toro, Panama/ 1996		FJ171792	FJ177544	FJ195659
	Aarti - Pan3	WA - Bocas del Toro, Panama/ 1996		FJ171792	FJ177544	FJ195660
	Aarti - SCa1	WA - Cape Island, South Carolina, USA; 12-13 m	USC S713	FJ171795	FJ177545	
	Aarti - SCa2	WA - Cape Island, South Carolina, USA; 12-13 m	USC S713	FJ171796	FJ177546	
A. bispinosus	Abisp - Her1	EM - Hersonissos, Crete, Greece; 3-5 m/ Mar 2006		FJ171743	FJ177568	
	Abisp - Her2	EM - Hersonissos, Crete, Greece; 3-5 m/ Mar 2006		FJ171743	FJ177568	
	Abisp - Her3	EM - Hersonissos, Crete, Greece; 3-5 m/ Mar 2006		FJ171743	FJ177568	
	Abisp - Cre1	EM - Cres, Croatia; 20 m/ Oct 2002		FJ171742	FJ177567	FJ195682
	Abisp - Cre2	EM - Cres, Croatia; 20 m/ Oct 2002		FJ171739	FJ177564	FJ195676
	Abisp - Cre3	EM - Cres, Croatia; 20 m/ Oct 2002		FJ171740	FJ177565	
	Abisp - Sar1	WM - Costa Colostrai, Sardinia; 17 m/ Aug 2002		FJ171745	FJ177570	FJ195733
	Abisp - Sar2	WM - Costa Colostrai, Sardinia; 17 m/ Aug 2002		FJ171741	FJ177566	FJ195680
	Abisp - Sar3	WM - Costa Colostrai, Sardinia; 17 m/ Aug 2002		FJ171742	FJ177567	FJ195682
	Abisp - Cor1	WM - St. Florent, Corsica; 1.5 m/ Apr 2005		FJ171744	FJ177569	
A. cingulatus	Acing - Pan1	WA - Isla Escuda de Veraguas, Panama; 42-39 m/ Aug 2004		FJ171802	FJ177552	FJ195664
A. comptus	Acomp - Flo1	WA - Gulf of Mexico, off St. Petersburg, Florida; 116 m/ Nov 2004	UF 3249	FJ171788	FJ177551	
	Acomp - Flo2	WA - Gulf of Mexico, off St. Petersburg, Florida; 116 m/ Nov 2004	UF 3249	FJ171788	FJ177551	
	Acomp - Flo3	WA- Off Sanibel Island, Florida, USA; 184 m/ 2001	UF 490	FJ171790	FJ177549	
A. duplicatus	Adupl - SCa1	WA - Cape Island, South Carolina, USA; 10 m	USC S1097	FJ171800	FJ177539	
	Adupl - SCa2	WA - Cape Island, South Carolina, USA; 10 m	USC S1098	FJ171801	FJ177540	
	Adupl - Bim1	WA - Bimini, Bahamas; 0 m/ Feb 2003		FJ171797	FJ177537	FJ195717
	Adupl - Bim2	WA - Bimini, Bahamas; 0 m/ Feb 2003		FJ171798	FJ177538	FJ195721
	Adupl - Flo1	WA - Sanibel Island, Florida, USA, 15-16m, 2001	UF 115, *A. forbesi*	FJ171799	FJ177542	
A. erinaceus	Aerin - Pan1	EP - Isla Montuosa, Golf of Chiriqui, Panama; 42.6 m/ May 2004		FJ171778	FJ177556	FJ195656
A. granulatus	Agran - Aus1	IP - Cobourg Peninsula, Northern Territory, Australia; 13 m/ Sep 1985	SI NMNH E38949	FJ171825	FJ177620	FJ195685
A. indicus	Aindi - Pak1	IO - Clifton, Karachi, Pakistan/ 2005		FJ171820	FJ177617	FJ195690
	Aindi - Pak2	IO - Clifton, Karachi, Pakistan/ 2005		FJ171820	FJ177617	FJ195691
	Aindi - Bru1	IP - Serasa, Brunei/ Dec 2004		FJ171821	FJ177618	FJ195726
	Aindi - Tha1	IP - Phuket, Thailand/ Apr 1997	PMBC 19223, *A. monacanthus*	FJ171819	FJ177619	FJ195692
A. irregularis	Airre - Sar1	WM - Costa Colostrai, Sardinia; 0-3 0m/ Jun 2001		FJ171749	FJ177574	FJ195677
	Airre - Sar2	WM - Costa Colostrai, Sardinia; 0-30m/ Jun 2001		FJ171751	FJ177576	FJ195693
	Airre - Sar3	WM - Costa Colostrai, Sardinia; 0-30m/ Aug 2002		FJ171750	FJ177575	FJ195678
	Airre - Nor1	EA - Irish Sea/ Jun-Oct 2001	Genbank		AY652501	
	Airre - Far1	EA - off Faro, Portugal; 530-540 m/ May 2005		FJ171758	FJ177583	FJ195665
	Airre - Far2	EA - off Faro, Portugal; 290 m/ May 2005		FJ171758	FJ177584	FJ195666
	Airre - Far3	EA - off Faro, Portugal; 106-128 m/ May 2005		FJ171758	FJ177584	FJ195667
	Airre - Kav1	EM - Kavala, Greece; > 50 m/ Mar 2006		FJ171752	FJ177577	FJ195696
	Airre - Kav2	EM - Kavala, Greece; > 50 m/ Mar 2006		FJ171753	FJ177578	FJ195697
	Airre - Kav3	EM - Kavala, Greece; > 50 m/ Mar 2006		FJ171754	FJ177579	FJ195698
	Airre - Mad1	EA - Quinta do Lorde, Madeira; 15-25m/ Oct 2005		FJ171755	FJ177580	FJ195732
A. javanicus	Ajava - Bru1	IP - Serasa, Brunei/ Dec 2004		FJ171824	FJ177616	FJ195725

A. jonstoni	Ajons - Tso1	EM - Tsoutsouras, Greece; 4 m/ Mar 2006		FJ171761	FJ177587	FJ195701
	Ajons - Tso2	EM - Tsoutsouras, Greece; 4 m/ Mar 2006		FJ171763	FJ177589	FJ195701
	Ajons - Tso3	EM - Tsoutsouras, Greece; 4 m/ Mar 2006		FJ171764	FJ177589	FJ195701
	Ajons - Sar1	WM - Costa Colostrai, Sardinia/ Aug 2000		FJ171762	FJ177588	FJ195703
	Ajons - Sar2	WM - Costa Colostrai, Sardinia; 1-30 m/ Aug 2002		FJ171762	FJ177590	FJ195703
	Ajons - Sar3	WM - Costa Colostrai, Sardinia, 1-30 m / Jun 2001		FJ171759	FJ177585	FJ195699
	Ajons - Sar4	WM - Costa Colostrai, Sardinia, 1-30 m / Jun 2001		FJ171760	FJ177586	FJ195700
A. latespinosus	Alate - Jap1	WP - Toyama Bay, Japan/ 2004		FJ171829	FJ177623	FJ195722
A. marginatus	Amarg - Pan1	WA - Colon, Panama; 10 m/ 2004		FJ171803	FJ177553	FJ195662
	Amarg - Pan2	WA - Colon, Panama; 10 m/ 2004		FJ171804	FJ177554	FJ195663
	Amarg - Col1	WA - Santa Marta, Colombia/ 2003	INV EQU02562	FJ171805	FJ177555	FJ195715
A. monacanthus	Amona - Aus1	IP - Torres Strait, Queensland, Australia; 0 m/ Jun 1979	SI NMNH E35296	FJ171826	FJ177621	FJ195735
A. nitidus	Aniti - Col1	WA - Santa Marta, Colombia; 153 m/ 2001	INV EQU01841	FJ171789	FJ177550	FJ195716
A. oerstedii	Aoers - Pan1	WA - San Blas, Panama/ Feb 2002		FJ171782	FJ177561	FJ195655
A. platyacanthus	Aplat - Cyp1	EM - Cap Greco, Cyprus; 10 m/ Oct 2004		FJ171748	FJ177573	FJ195724
	Aplat - Cyp2	EM - De Capo Bay, Cyprus; 3 m/ Oct 2004		FJ171748	FJ177573	FJ195724
	Aplat - Cyp3	EM - Cap Greco, Cyprus; 10 m/ Oct 2004		FJ171748	FJ177573	FJ195724
	Aplat - Sar1	WM - Sardinia; 1-30m/ 2002		FJ171747	FJ177572	FJ195684
	Aplat - Sar2	WM - Sardinia; 1-30m / 2002		FJ171747	FJ177572	FJ195684
	Aplat - Sar3	WM - Cannigione, Sardinia, 1-30m/ Sep 2003		FJ171747	FJ177572	FJ195684
	Aplat - LaH1	WM - La Herradura, Spain; 5 m/ Aug 2005		FJ171746	FJ177571	
	Aplat - Tou1	WM - Toulon, France/ May 2000		FJ171747	FJ177572	
A. polyacanthus	Apoly - Mau1	CP - Maui, Hawaiian Islands; 10 m/ Apr 2005	LACM 2005-64.1	FJ171807	FJ177603	
	Apoly - Mau2	CP - Kanaio, Maui, Hawaii; 10 m/ Feb 2006		FJ171808	FJ177602	FJ195694
	Apoly - Sey1	IO - Picard Island, Aldabra Islands, Seychelles; reef flat/ Mar 1987	SI NMNH E35107	FJ171810	FJ177607	FJ195686
	Apoly - Sey2	IO - Picard Island, Aldabra Islands, Seychelles; reef flat/ Mar 1987	SI NMNH E35107	FJ171811	FJ177608	FJ195688
	Apoly - Fij1	SP - Bligh Water, Fiji; 143-173 m/ Aug 1998	MNHN EcAh 4734	FJ171827		
	Apoly - NZ1	SP - Auckland, New Zealand/ 2003		FJ171813	FJ177610	FJ195711
	Apoly - NZ2	SP - Auckland, New Zealand/ 2003		FJ171814	FJ177611	FJ195712
	Apoly - Jap1	WP - Toyama Bay, Japan/ Nov 2003		FJ171806	FJ177605	FJ195719
	Apoly - Jap2	WP - Toyama Bay, Japan/ Nov 2003		FJ171806	FJ177606	FJ195720
	Apoly - Phi1	IP - Aligbay, Mindanao, Philippines; 1.5 m/ May 1979	SI NMNH E48900	FJ171812	FJ177609	FJ195736
	Apoly - Dub1	IO - Dugass Beach, Dubai, U.E.A./ Feb 1981	SI NMNH E35065	FJ195652	FJ177633	
A. regalis	Arega - CRi1	EA - Coco, Guanacaste, Costa Rica; 4 m/ 1933	LACM 1933-123	FJ171780	FJ177558	FJ195653
A. scoparius	Ascop - Jap1	WP - Toyama Bay, Japan/ Nov 2003		FJ171818	FJ177613	FJ195718
A. siderealis	Aside - Pan1	EP - Coiba Island, Panama; 58 m/ May 2004		FJ171784	FJ177562	FJ195657
A. sp. 1	Asp1 - Fij1	SP - Malolo, Viti Levu, Fiji; 39 m/ Oct 1998	MNHN EcAh 4730	FJ171817		
A. sp. 2	Asp2 - Fij1	SP - SE Viti Levu, Fiji; 244-252 m/ Aug 1998	MNHN EcAh 4725	FJ171836	FJ177627	FJ195702
A. sp. 3	Asp3 - Fij1	SP - NW Taveuni Island, Fiji; 327-420 m/ Mar 1999	MNHN EcAh 4729	FJ171830	FJ177624	FJ195704
	Asp3 - Fij2	SP - off Suva, Fiji; 478-500m/ Mar 1999	MNHN EcAh 4735	FJ171832	FJ177625	FJ195705
	Asp3 - Fij3	SP - Bligh Water, N Viti Levu, Fiji; 471-475 m/ Aug 1998	MNHN EcAh 4759	FJ171831		FJ195710
A. sp. 4; (possibly *A. tasmanicus* or *A. eremicus*)	Asp4 - Fij1	SP - S Namenalala, Fiji; 364-369 m/ Mar 1999	MNHN EcAh 4738	FJ171833	FJ177626	FJ195707
	Asp4 - Fij2	SP - S Namenalala, Fiji; 364-369 m; Mar 1999	MNHN EcAh 4739	FJ171834	FJ177626	
A. sp. 5	Asp5 - Ton1	SP - SW Tongatapu, Tonga; 319-333 m/ Jun 2000	MNHN EcAh 4741	FJ171835		FJ195708
A. sp. 6	Asp6 - Phi1	IP - Leyte Island, Philippines; 76 m/ Nov 1979	SI NMNH E53739	FJ171828	FJ177622	FJ195737
A. sp. 7	Asp7 - Aus1	SP - Pallarenda Beach, Townsville, Australia; 0 m/ May 2004		FJ171823	FJ177615	FJ195728
A. spinulosus	Aspin - Cre1	EM - Cres, Croatia; 5 m/ Oct 2002		FJ171756	FJ177581	FJ195706
	Aspin - Cor1	WM - St. Florent, Corsica; 1.5m/ Apr 2005		FJ171757	FJ177582	FJ195723
A. triseriatus	Atris - Oah1	EP - Kailua, Oahu, Hawaiian Islands; 22m/ 1980	BM 1980.536	FJ171809	FJ177604	
A. vappa	Avapp - Bru1	IP - Serasa, Brunei/ Dec 2004		FJ171816	FJ177612	FJ195727
A. verilli	Averr - Pan1	EP - Panama/ 1996?		FJ171779	FJ177557	FJ195661
	Averr - Cal1	EP - Monterrey Bay, California, USA/ May 1996	Calacad 105628	FJ171783	FJ177560	FJ195713
	Averr - Cal2	EP - Point Loma, San Diego, USA; 220 m/ Nov 2002	SIO BIC E3481	FJ171781	FJ177559	FJ195730
	Averr - Cal3	EP - Point Loma, San Diego, USA; 220 m/ Nov 2002	SIO BIC E3481	FJ171781	FJ177559	FJ195731
A. zebra	Azebr - PNG1	SP - Deboin Mission, SE Papua New Guinea,/ Jun 1979	SI NMNH E50681	FJ171822	FJ177614	FJ195689
Tethyaster sp. 1	Tethy - Ton1	SP - Eua, Tonga; 463-464m/ Jun 2000	MNHN EcAh 4758	FJ171840		FJ195709
Tethyaster sp. 2	Tethy - Naz1	EP - Nazca submarine ridge; 230-280 m/ May 1987		FJ171841	FJ177630	FJ195734
Ctenophoraster sp.	Cphor - Mar1	SP - Fatu Hiva, Marquesas Islands; 85-130m/ Sep 1997	MNHN EcAh 4749	FJ171837		
Ctenopleura sp.	Cpleu - Fij1	SP - Bligh Water, N Viti Levu, Fiji; 143-173 m/ August 1998	MNHN EcAh 4732	FJ171843	FJ177632	

Pseudarchaster parelii	Ppare - Pan1	WA - San Blas, Panama/ Feb 2003		FJ171838	FJ177628	FJ195738
Tethyaster subinermis	Tsubi - Far1	EA - off Portimão, Faro, Portugal; 120-131m/ May 2005		FJ171839	FJ177629	FJ195672
Thrissacanthias penicillatas	Tpeni - SDi1	EP - off San Diego, California, USA; 1215 m/ Oct 2005	SIO BIC E3857	FJ171842	FJ177631	FJ195673
Total	**125**					

EA = East Atlantic; WM = West Mediterranean; EM = East Mediterranean; WA = West Atlantic; EP = East Pacific; CP = Central Pacific; SP = South Pacific; WP = West Pacific; IP = Indo-Pacific; IO = Indian Ocean. BM = Bishop Museum, Honolulu, Hawaii; Calacad = California Academy of Science; INV = INVEMAR, Instituto de Investigaciones Marinas y Costeras, Colombia; LACM = Los Angeles County Museum; MNHN = Musée nationale de la histoire naturelle, Paris, Echinoderm Collection; PMBC = Phuket Marine Biological Center; SI NMNH = Smithsonian National Museum of Natural History, Invertebrate Collection; SIO BIC = Scripps Institution of Oceanography - Benthic Invertebrate Collection; UF = Florida Museum of Natural History; UNAM = Universidad Nacional Autonoma de Mexico, Colleccion Nacional de Equinodermos; USC = University of South Carolina.

Table 2: Primers used for PCR-amplification of three mitochondrial DNA regions in *Astropecten* species.

Region	ID	Direction	Sequence 5'-3'	Reference
12S	12Saf*	forward	CTT AGC AAC CGA TTT GGT CCT AGT CC	this study
	12Sar	reverse	GCT GGT AAG GTT TTT CGT GGG TTA TCG	this study
	12Sa2r *	reverse	CCG CCA AGT CCT TTG AG	this study
16S	16Sbr*	forward	CCG GTC T(C/G)A (GA/AC)T CAG ATC ACG	Palumbi *et al.* 1996
	16Sar *	reverse	CGC CTG TTT ACC (A/T)AA AAC AT	Palumbi *et al.* 1996
	16Sa3r*	reverse	GTT AAA CGG CCG CGG TAT TTT GAC CG	this study
COI	cHCOf *	forward	(TGA/GAT) TTT TTG GTC ACC CTG AAG TTT A	Folmer *et al.* 1994
	AstroCOImf	forward	TAC TAT GTT GTA GCA CAC TT	this study
	ECOIa	forward	ACC ATG CAA CTA AGA CGA TGA	Smith *et al.* 1993
	AstroCOI2r*	reverse	TCT GAG TAT CGT CGT GGC ATT CC	this study
	COIe*	reverse	CCA GAG AAG AGG GGA AAC CAG TG	Palumbi 1996
	AstroCOImr*	reverse	AAG TGT GCT ACA ACA TAG TA	this study
	ECOIb*	reverse	GGT AGT CTG AGT ATC GTC G(AT)G	Knott *et al.* 2000

* = primers used for cycle sequencing reaction

Table 3: Genetic distances between specimens of *A. irregularis*. Above diagonal: uncorrected genetic distances; below diagonal: genetic distances corrected using parameters estimated by Modeltest.

	Far1	**Far2, Far3**	**Kav1**	**Kav2**	**Kav3**	**Mad1**	**Nor1**	**Sar1**	**Sar2**	**Sar3**
Far1	*	0.002	0.082	0.078	0.079	0.084	0.053	0.062	0.079	0.080
Far2, Far3	0.073	*	0.082	0.078	0.079	0.084	0.053	0.061	0.079	0.079
Kav1	0.106	0.106	*	0.005	0.004	0.005	0.019	0.015	0.008	0.035
Kav2	0.100	0.100	0.004	*	0.001	0.005	0.016	0.012	0.006	0.033
Kav3	0.101	0.101	0.004	0.001	*	0.006	0.016	0.013	0.006	0.034
Mad1	0.111	0.111	0.005	0.005	0.005	*	0.016	0.014	0.011	0.038
Nor1	0.062	0.062	0.019	0.017	0.017	0.017	*	0.017	0.017	0.002
Sar1	0.075	0.073	0.016	0.012	0.013	0.014	0.017	*	0.014	0.021
Sar2	0.101	0.101	0.008	0.006	0.005	0.011	0.018	0.014	*	0.037
Sar3	0.104	0.101	0.038	0.036	0.037	0.042	0.002	0.022	0.041	*

Figure 1: Phylogeny of the genus *Astropecten* as suggested by Döderlein (1917) presenting the relationships of species and species groups relevant to this study.

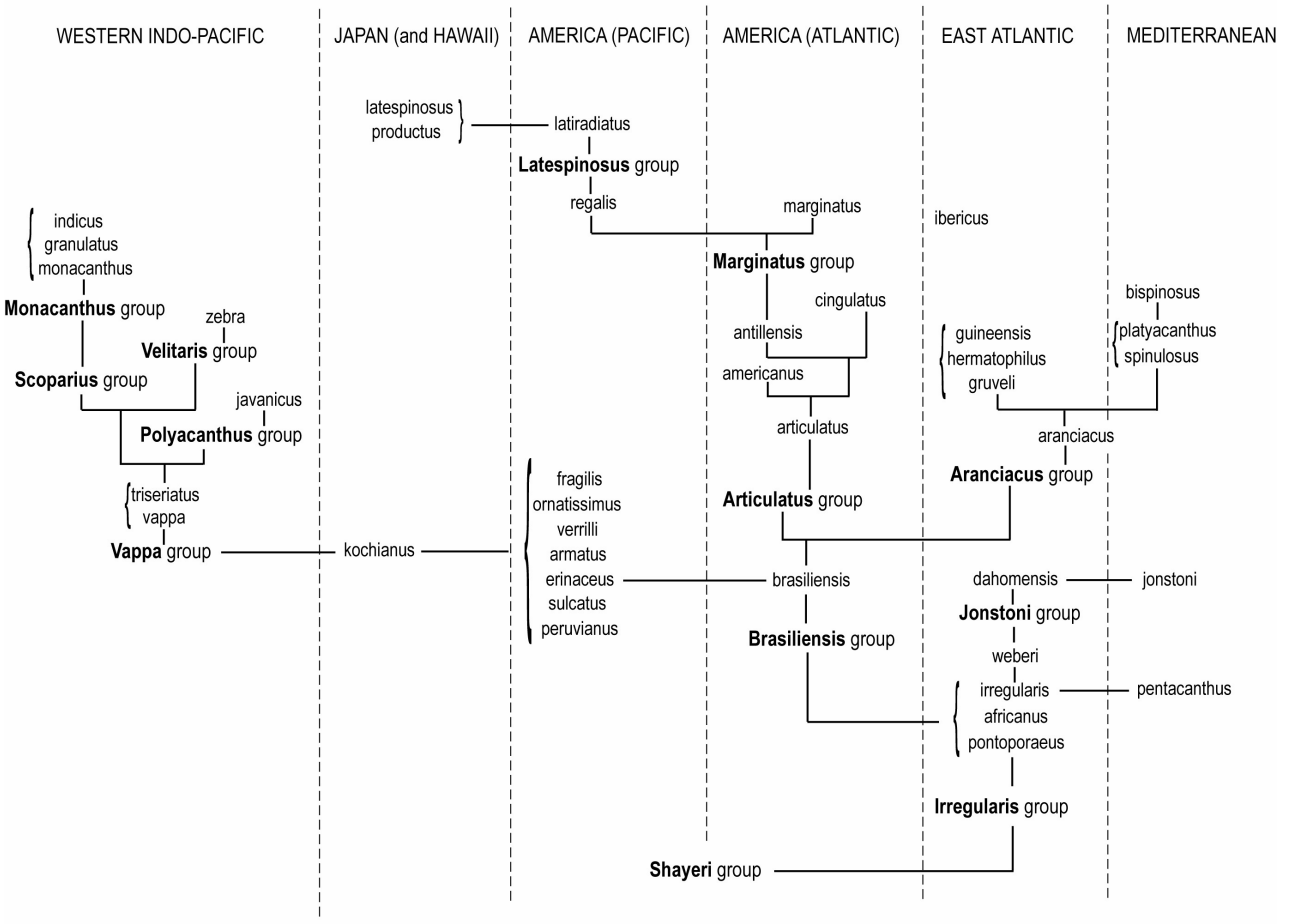
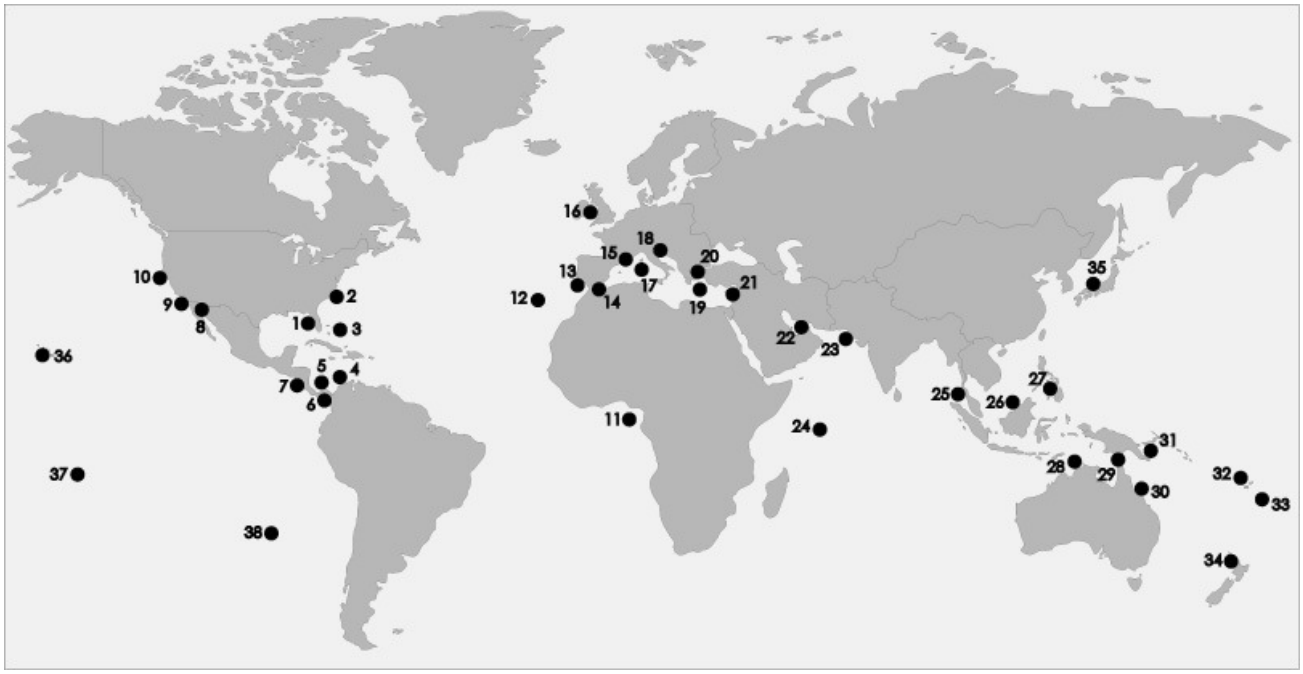


Figure 2: Collection sites of *Astropecten* specimens and outgroup taxa.



1. Florida: Sanibel Island and St. Petersburg; 2. South Carolina: Cape Island,; 3. Bahamas: Bimini,; 4. Colombia: St. Marta and Arboletes; 5. Atlantic Panama: Bocas del Toro, San Blas and Colon; 6. Pacific Panama: Gulf of Chiriqui, Bay of Panama; 7. Costa Rica: Guanacaste; 8. Mexico: Puerto Peñasco; 9. California: San Diego; 10. California: Monterrey Bay; 11. São Tomé; 12. Madeira; 13. Portugal: Faro; 14. Spain: La Herradura; 15. France: Toulon; 16. Irish Sea; 17. West Mediterranean: Sardinia and Corsica; 18. Croatia: Cres; 19. Greece: Crete; 20. Greece: Kavala; 21. Greece: Cyprus; 22. United Arab Emirates: Dubai; 23. Pakistan: Karachi; 24. Seychelles; 25. Thailand: Phuket; 26. Borneo: Brunei; 27. Philippines: Mindanao and Leyte; 28. Australia: Cobourgh Peninsula; 29. Australia: Torres Strait; 30. Australia: Townsville; 31. SE Papua New Guinea; 32. Fiji; 33. Tonga; 34. Auckland, New Zealand; 35. Toyama Bay, Japan; 36. Hawaii: Maui and. Oahu; 37. Marquesas Islands; 38. Nazca submarine rigde

Figure 3: Total evidence molecular phylogeny of 118 specimens of *Astropecten* and seven outgroup taxa based on combined 12S, 16S and COI sequences of the mitochondrial DNA. (a) Maximum parsimony strict consensus tree of 800 most parsimonious trees (*length* = 3564). Above nodes: bootstrap support values above 50% from 1'000 replicates. Below nodes: decay indices. (b) Bayesian inference tree resulting from 8'000 trees based on the GTR+I+Γ model. Posterior probabilities are indicated above nodes. See Table 1 for locality abbreviations. EP = East Pacific, WA = West Atlantic; CP = Central Pacific; WP = West Pacific; IP = Indo-Pacific; IO = Indian Ocean; EA = East Atlantic, EM = East Mediterranean, WM = West Mediterranean

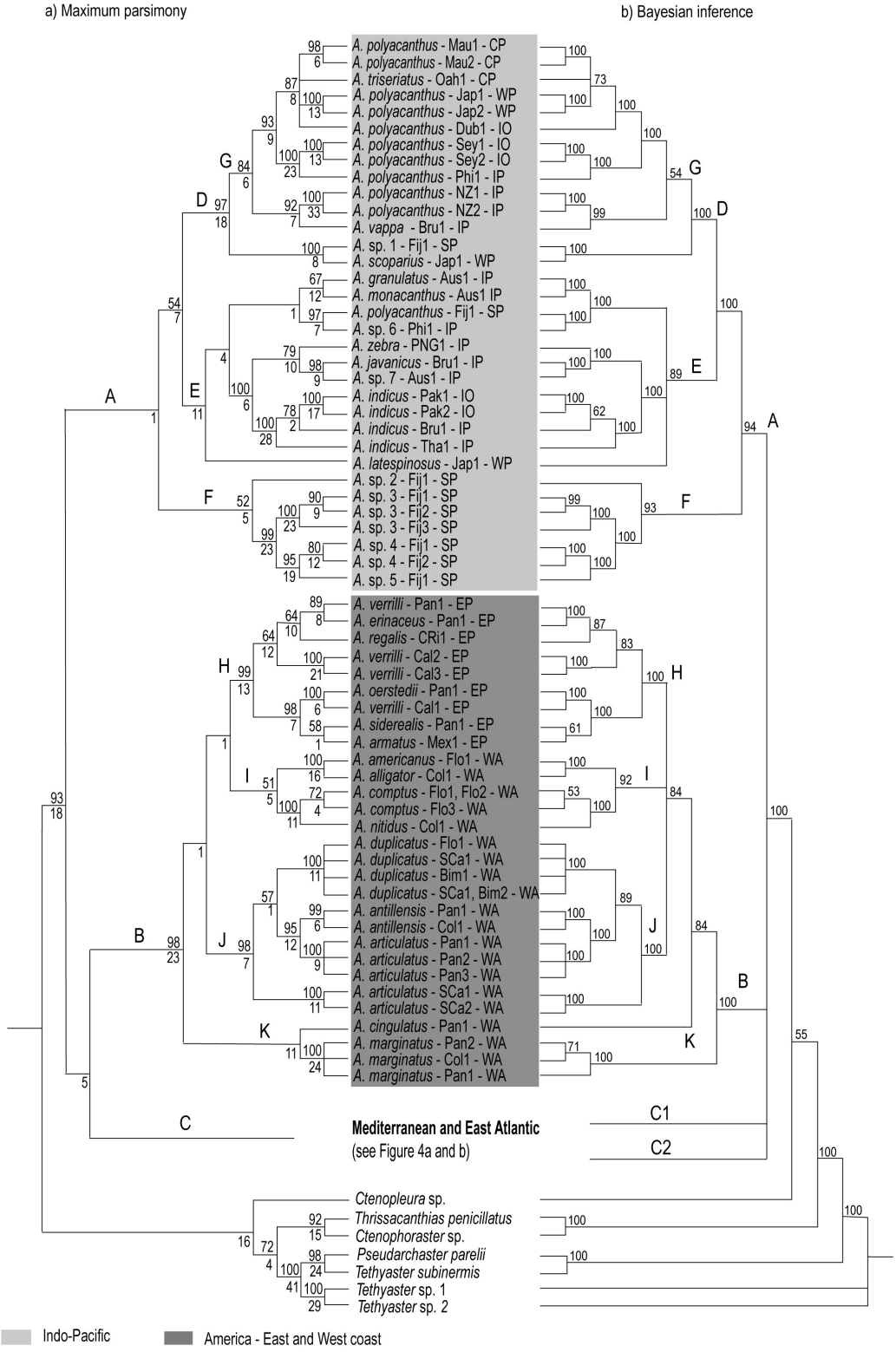


Figure 4: Mediterranean and East Atlantic (see caption of Figure 3 for further explanations)

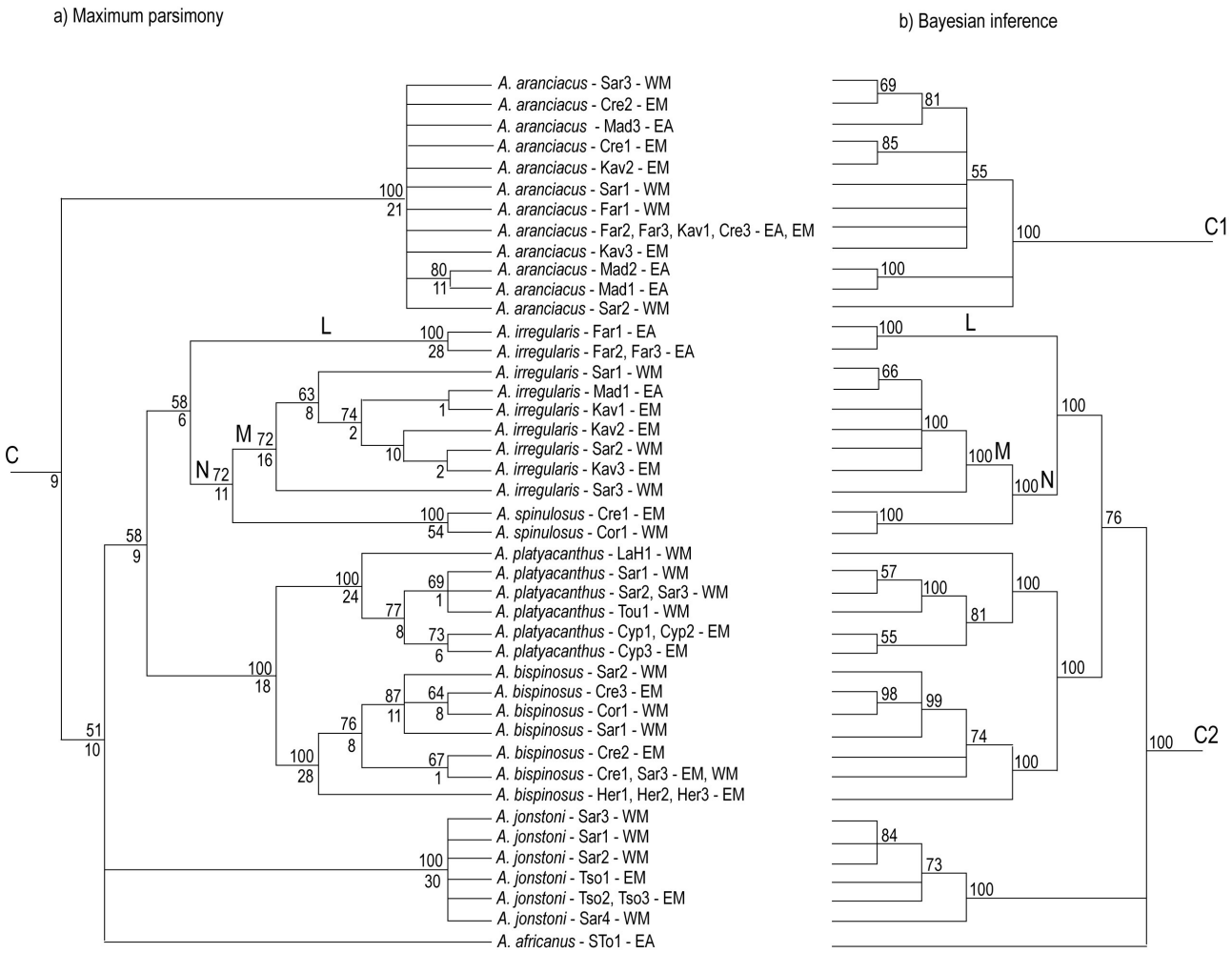
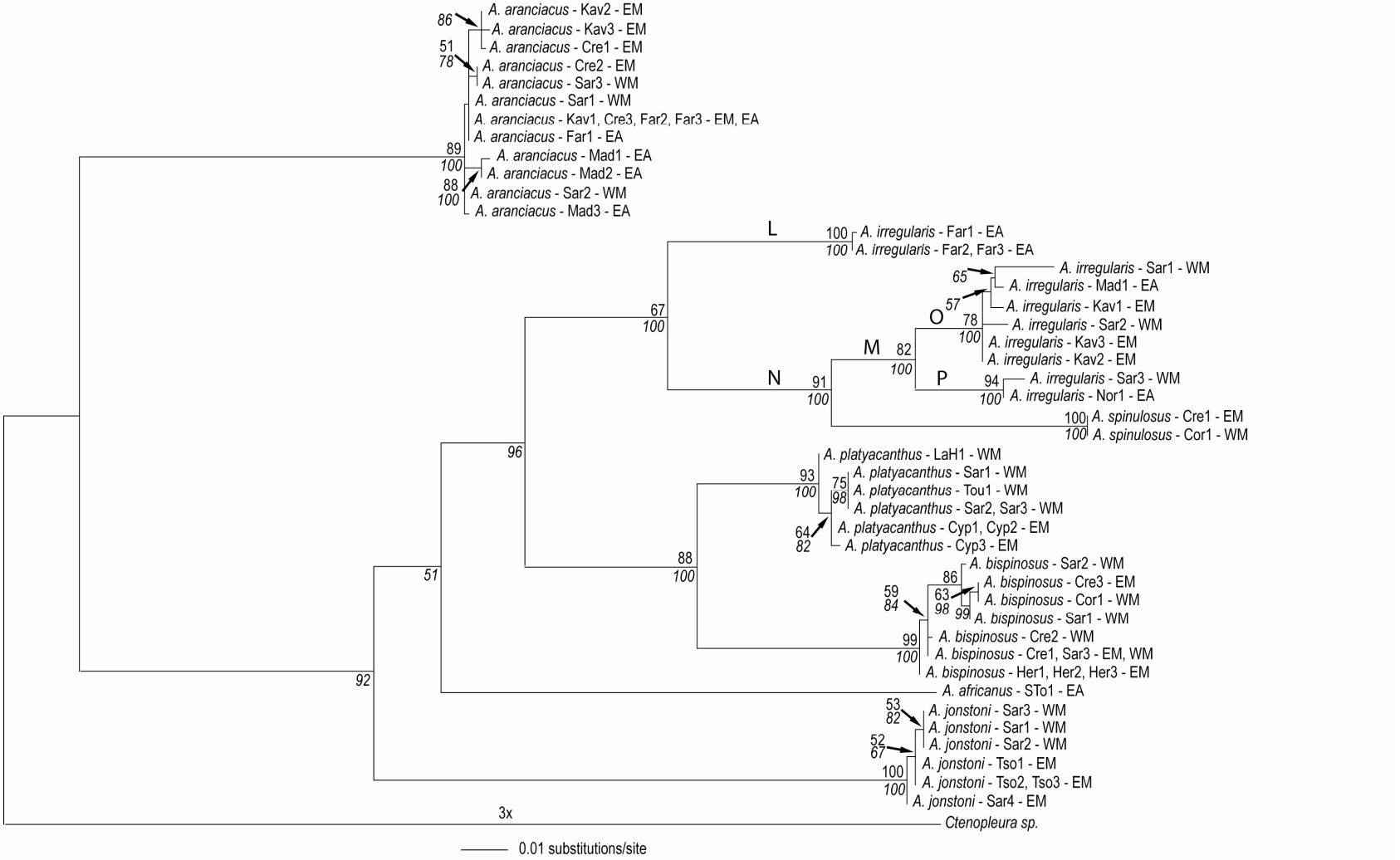


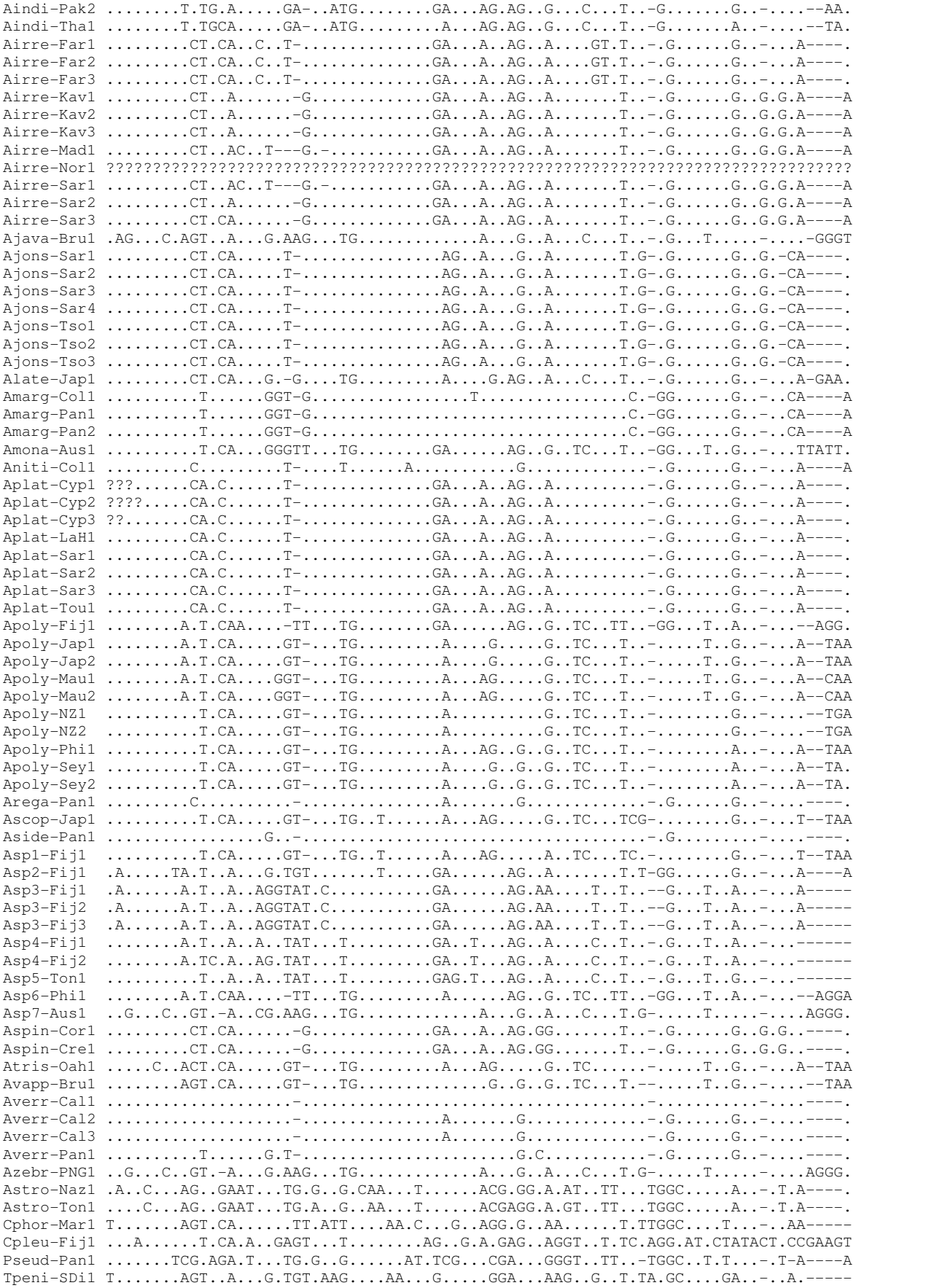
Figure 5: Maximum likelihood tree (-lnL = 5480.33; estimated proportion of invariable sites = 0.487; estimated values of gamma shape parameter = 0.444) of 54 Mediterranean and East Atlantic *Astropecten* specimens and one outgroup specimen (*Ctenophoraster*). ML tree was obtained by heuristic search using the TBR branch swapping option and the TVM+I+Γ substitution model. Above nodes: bootstrap support values > 50% from 1000 replicates. Below nodes in italic: posterior probabilities from Bayesian analysis > 50%.

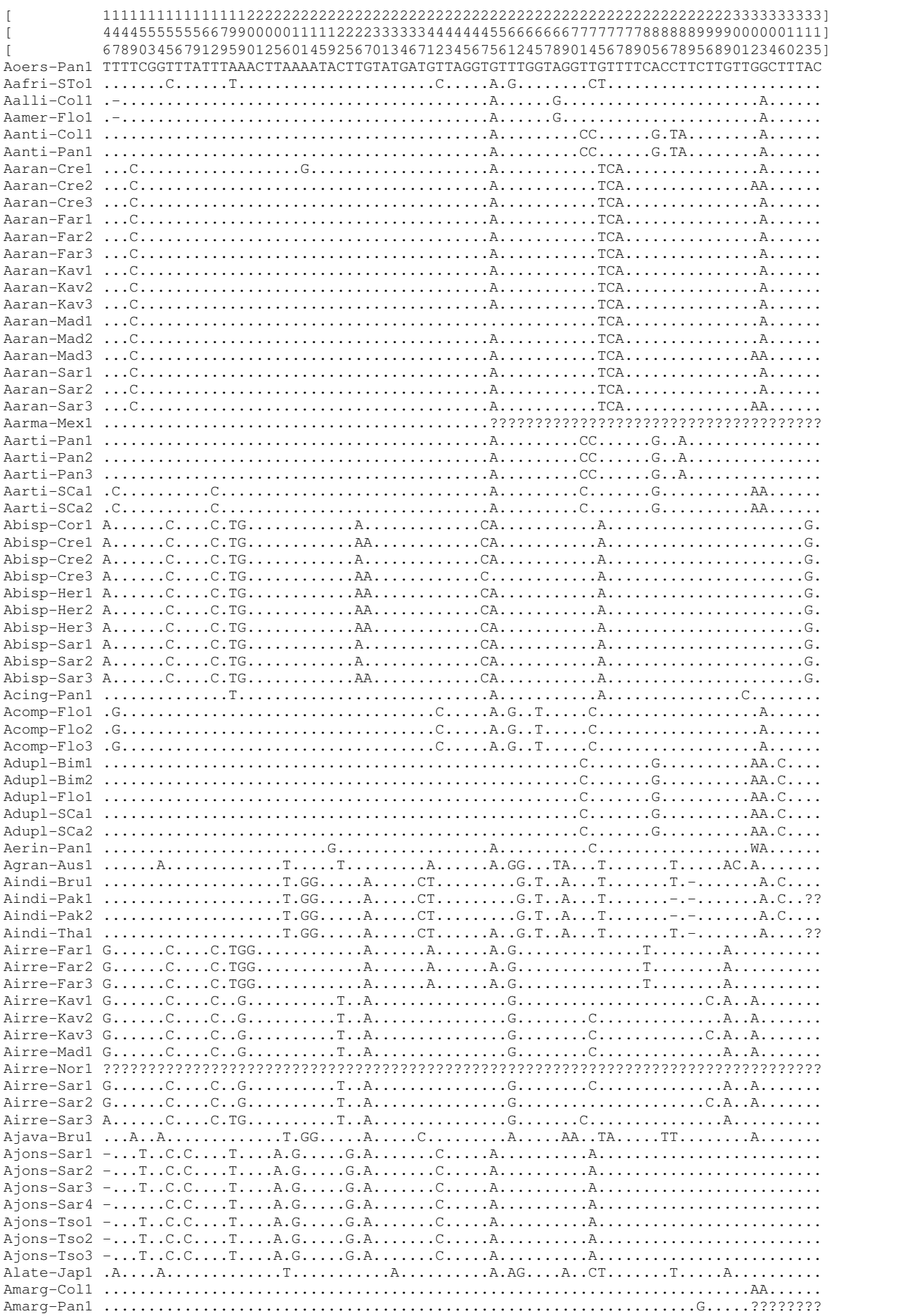


Appendix: Sequence alignment of 118 samples of *Astropecten* and 7 outgroup taxa showing the reference sequence (Aoers-Pan1) and the variable sites.

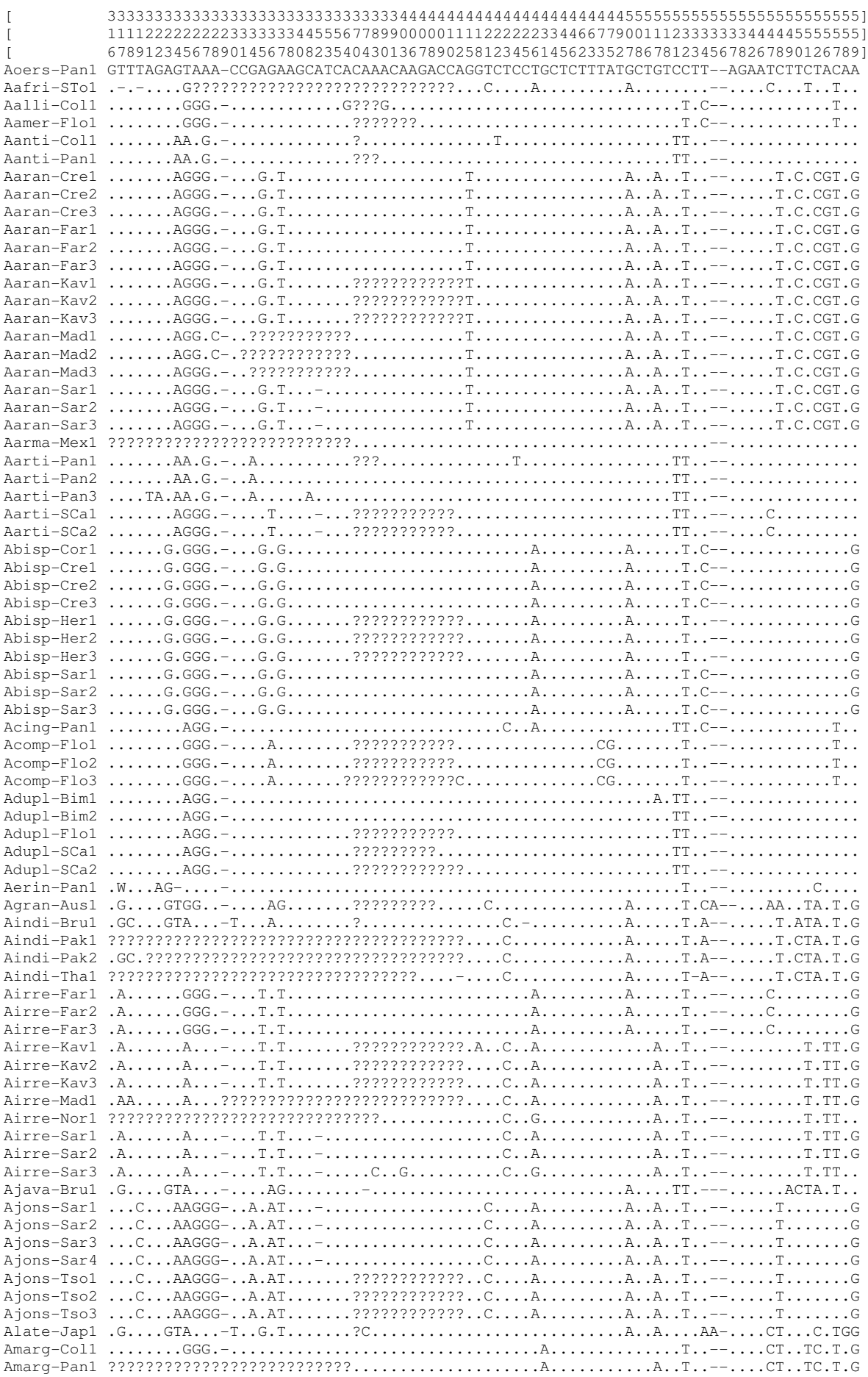
```
Aoers-Pan1
GGCTCATATA GTGGGGTATC TAAT-CCCAG TTTATTTTTG AGCTTTGGTG TCTTGGGTAA C-ATTGTAAT TAGTTTACTG TAGAAGCTTG
TTAATACTGA AGTTTTTTTC TACCCTAAA- AATCACCTTT ATACC-GATG ----GTTTTC TTGGTTTGAG TTTTATGTAT AACCGCGGTG
GCTGGCACAT AATTAACCAA CTCTTATATA ACTATATCAA GCTTTTGCTT ATTGTATAAT GTTAGGTACT GCAGGTGTGG TT-GGCTAGG
TGTTGTTTTC TGATACCTTA CTCACTGTGT TGGCTTGCAT GTATCGTTTT AGAGTAAA-C CAAGAGAAGG ACTAGGACCA AAT-CGGTTG
CTAAGAGGTC TGAACTCAGA TCACGTAGGA CTTTAATGGT CGAACAGACC AACTGCTGCA TCTCCTGGAA GTCCCGATCC AACATCGAGG
T-CGCAAACT TTTTTATCAA TATGAACT-C TCCAAAAAAA TTACGCTGTT ATCCCTGCGG TAACTTGTTC CTTTGATCAC TCTT--AGTG
GATCAATCTT CTAAAACAAT AAAT-AGCGG AGGGTTTGTT TTCTCCGTGG TTGCCCCAAC CAAAACAACT AA-TTAGCAG TA-AACTTTT
AATT-CCGAG TAAAA-TCAA AATTAATAAC TCTAATTC-A TAGCTTAAGC TCGACAGGG- TCTTCTCGTC CTACGACTTT ATCCCCGCCT
CTTCACGAGG ATGATAAGTT CAAGATTCTA AAAGAGAGAC AGTTAAAC-T TCCGTCTTGC CATTCATACC AGCCTCCATT TAAAAGGCAA
ATGATTATGC TACCTTTGCA CGGTCAAAAT ACCGCGGCCG TTTATGATCT CCCACGTAAT AGCACACTAC TCAGGAAAGG CCGAACCTTT
TGGTTATCTA GGAATGGTTT ACGCAATAGT CTCAATAGGA ATATTAGGAT TCCTAGTTTG AGCACACCAC ATGTTTACAG TCGGGATGGA
CGTAGACACC CGAGCATACT TCACCGCTGC AACGATGATA ATTGCTGTCC CAACTGGAAT AAAGGTATTT AGCTGAATGG CCACCCTTCA
AGGAAGAAAC TTACGATGAG ACACCCCACT ACTTTGAGCC CTAGGATTCG TGTTCCTCTT CACAATTGGA GGACTTACTG GGGTGATTCT
AGCAAAATCA TCAATTGACG TTGTTCTTCA CGATACCTAC TACGTTGTAG CACACTTTCA CTACGTTCTT TCAATGGGGG CAGTATTCGC
AATATTTGCA GGGTTCACTC ATT?GGTTCC CGCTTTTCTC AGGACTCTCT TTTCACCCCC TCTGAGGAAA GATACACTTT ATTCTTATGT
TTATAGGGGT TAACCTAACT TTTTTCCCAC AACACTTCCT


[                                                  11111111111111111111111111111111111]
[            13334455555555566666666677777778888889999000000111111112222222233333333444444]
[          12616793712345678901234568901345670234570469034568012345690123467912356789012345]
Aoers-Pan1 GGAGTTTCGTCTTGGGTAAC-ATTGAATTGTTTAGAGAACGAAGATTTTTCTACCCTA-AATCCCTATAC-GATG----G
Aafri-STo1 ..........T.CA...T.T-....G..........GA......AG..A..T....T..-.G...T.G...-...A----A
Aalli-Col1 ....................T-....T..................G..............T.G......G..-...A----T
Aamer-Flo1 ....................T-....T..................G..............T.G......G..-...A----T
Aanti-Col1 ..........T......GGTC....G..............A...G............G-.G......G..-.GCA----.
Aanti-Pan1 ..........T......GGTC....G..............A...G............G-.G......G..-.GGA----.
Aaran-Cre1 .........CT.CA...GGT-....T.........GA......AG..A.......T..-.G......G..-....----A
Aaran-Cre2 .........CT.CA...GGT-....T.........GA......AG..A.......T..-.G......G..-....----A
Aaran-Cre3 .........CT.CA...GGT-....T.........GA......AG..A.......T..-.G......G..-....----A
Aaran-Far1 .........CT.CA...GGT-....T.........GA......AG..A.......T..-.G......G..-....----A
Aaran-Far2 .........CT.CA...GGT-....T.........GA......AG..A.......T..-.G......G..-....----A
Aaran-Far3 .........CT.CA...GGT-....T.........GA......AG..A.......T..-.G......G..-....----A
Aaran-Kav1 .........CT.CA...GGT-....T.........GA......AG..A.......T..-.G......G..-....----A
Aaran-Kav2 .........CT.CA...GGT-....T.........GA......AG..A.......T..-.G......G..-....----A
Aaran-Kav3 .........CT.CA...GGT-....T.........GA......AG..A.......T..-.G......G..-....----A
Aaran-Mad1 .........CT.CA...GGT-....T.........GA......AG..A.......T..-.G......G..-....----A
Aaran-Mad2 .........CT.CA...GGT-....T.........GA......AG..A.......T..-.G......G..-....----A
Aaran-Mad3 .........CT.CA...GGT-....T.........GA......AG..A.......T..-.G......G..-....----A
Aaran-Sar1 .........CT.CA...GGT-....T.........GA......AG..A.......T..-.G......G..-....----A
Aaran-Sar2 .........CT.CA...GGT-....T.........GA......AG..A.......T..-.G......G..-....----A
Aaran-Sar3 .........CT.CA...GGT-....T.........GA......AG..A.......T..-.G......G..-....----A
Aarma-Mex1 ..................G..-.....................................-.G.........-C...----.
Aarti-Pan1 ..................GGTC....G..................G............G-.G......G..-.GCA----.
Aarti-Pan2 ..................GGTC....G..................G............G-.G......G..-.GCA----.
Aarti-Pan3 ??................GGTC....G..................G............G-.G......G..-.GCA----.
Aarti-SCa1 ..................GGT-....G.........G........G............G-.G......G..-..C.----.
Aarti-SCa2 ..................GGT-....G..................G............G-.G......G..-..C.----.
Abisp-Cor1 .........CT.CA.....T-..............GA...A..AG..A.......T..-.G......G..-..CA----.
Abisp-Cre1 .........CT.CA.....T-..............GA...A..AG..A.......T..-.G......G..-..CA----.
Abisp-Cre2 .........CT.CA.....T-..............GA...A..AG..A.......T..-.G......G..-..CA----.
Abisp-Cre3 .........CT.CA.....T-..............GA...A..AG..A.......T..-.G......G..-..CA----.
Abisp-Her1 .........CT.CA.....T-..............GA...A..AG..A.......T..-.G......G..-..CA----.
Abisp-Her2 .........CT.CA.....T-..............GA...A..AG..A.......T..-.G......G..-..CA----.
Abisp-Her3 .........CT.CA.....T-..............GA...A..AG..A.......T..-.G......G..-..CA----.
Abisp-Sar1 .........CT.CA.....T-..............GA...A..AG..A.......T..-.G......G..-..CA----.
Abisp-Sar2 .........CT.CA.....T-..............GA...A..AG..A.......T..-.G......G..-..CA----.
Abisp-Sar3 .........CT.CA.....T-..............GA...A..AG..A.......T..-.G......G..-..CA----.
Acing-Pan1 ...A......T......G.T-G...G..................G..............-.G......G..-...A----A
Acomp-Flo1 .........C.........T-....T......A..........G..............-.G......G..-...A----A
Acomp-Flo2 .........C.........T-....T......A..........G..............-.G......G..-...A----A
Acomp-Flo3 .........C.........T-....T......A..........G..............-.G......G..-...A----A
Adupl-Bim1 ..........T.....CGGT-....G........A.........G..............-.G......G..-..GA----.
Adupl-Bim2 ..........T.....CGGT-....G........A.........G..............-.G......G..-..GA----.
Adupl-Flo1 ..........T.....CGGT-....G........A.........G..............-.G......G..-..GA----.
Adupl-SCa1 ..........T.....CGGT-....G........A.........G..............-.G......G..-..G.----.
Adupl-SCa2 ..........T.....CGGT-....G........A.........G..............-.G......G..-..GA----.
Aerin-Pan1 ...A......T......CGT-....G..................G.C............-.G......G..-.....----.
Agran-Aus1 ..........T.CA...GGGTT...TG........GA......AG..G..TC...T..-GG...T..G..-...TTATT.
Aindi-Bru1 ........T.T.GA.....GA-...TG........GA...AG.AG..G...C...T.G-G.......G..-....--AAA
Aindi-Pak1 ........T.TG.A.....GA-..ATG........GA...AG.AG..G...C...T..-G.......G..-....--AA.
```


Aindi-Pak2 ........T.TG.A.....GA-..ATG........GA...AG.AG..G...C...T..-G.......G..-....--AA.
Aindi-Tha1 ........T.TGCA.....GA-..ATG.........A...AG.AG..G...C...T..-G.......A..-....--TA.
Airre-Far1 .........CT.CA..C..T-..............GA...A..AG..A....GT.T..-.G......G..-...A----.
Airre-Far2 .........CT.CA..C..T-..............GA...A..AG..A....GT.T..-.G......G..-...A----.
Airre-Far3 .........CT.CA..C..T-..............GA...A..AG..A....GT.T..-.G......G..-...A----.
Airre-Kav1 .........CT..A......-G.............GA...A..AG..A.......T..-.G......G..G.G.A----A
Airre-Kav2 .........CT..A......-G.............GA...A..AG..A.......T..-.G......G..G.G.A----A
Airre-Kav3 .........CT..A......-G.............GA...A..AG..A.......T..-.G......G..G.G.A----A
Airre-Mad1 .........CT..AC..T---G.-...........GA...A..AG..A.......T..-.G......G..G.G.A----A
Airre-Nor1 ??
Airre-Sar1 .........CT..AC..T---G.-...........GA...A..AG..A.......T..-.G......G..G.G.A----A
Airre-Sar2 .........CT..A......-G.............GA...A..AG..A.......T..-.G......G..G.G.A----A
Airre-Sar3 .........CT.CA......-G.............GA...A..AG..A.......T..-.G......G..G.G.A----A
Ajava-Bru1 .AG...C.AGT..A...G.AAG...TG.............A...G..A...C...T..-.G...T.....-....-GGGT
Ajons-Sar1 .........CT.CA.....T-..............AG..A...G..A.......T.G-.G......G..G.-CA----.
Ajons-Sar2 .........CT.CA.....T-..............AG..A...G..A.......T.G-.G......G..G.-CA----.
Ajons-Sar3 .........CT.CA.....T-..............AG..A...G..A.......T.G-.G......G..G.-CA----.
Ajons-Sar4 .........CT.CA.....T-..............AG..A...G..A.......T.G-.G......G..G.-CA----.
Ajons-Tso1 .........CT.CA.....T-..............AG..A...G..A.......T.G-.G......G..G.-CA----.
Ajons-Tso2 .........CT.CA.....T-..............AG..A...G..A.......T.G-.G......G..G.-CA----.
Ajons-Tso3 .........CT.CA.....T-..............AG..A...G..A.......T.G-.G......G..G.-CA----.
Alate-Jap1 .........CT.CA...G.-G....TG.........A....G.AG..A...C...T..-.G......G..-...A-GAA.
Amarg-Col1 ..........T......GGT-G..................T..................C.-GG......G..-..CA----A
Amarg-Pan1 ..........T......GGT-G.................................C.-GG......G..-..CA----A
Amarg-Pan2 ..........T......GGT-G.................................C.-GG......G..-..CA----A
Amona-Aus1 ..........T.CA...GGGTT...TG........GA......AG..G..TC...T..-GG...T..G..-...TTATT.
Aniti-Col1 .........C.........T-....T......A...........G.............-.G......G..-...A----A
Aplat-Cyp1 ???......CA.C......T-..............GA...A..AG..A...........-.G......G..-...A----.
Aplat-Cyp2 ????.....CA.C......T-..............GA...A..AG..A...........-.G......G..-...A----.
Aplat-Cyp3 ??.......CA.C......T-..............GA...A..AG..A...........-.G......G..-...A----.
Aplat-LaH1 .........CA.C......T-..............GA...A..AG..A...........-.G......G..-...A----.
Aplat-Sar1 .........CA.C......T-..............GA...A..AG..A...........-.G......G..-...A----.
Aplat-Sar2 .........CA.C......T-..............GA...A..AG..A...........-.G......G..-...A----.
Aplat-Sar3 .........CA.C......T-..............GA...A..AG..A...........-.G......G..-...A----.
Aplat-Tou1 .........CA.C......T-..............GA...A..AG..A...........-.G......G..-...A----.
Apoly-Fij1 ........A.T.CAA....-TT...TG........GA......AG..G..TC..TT..-GG...T..A..-...--AGG.
Apoly-Jap1 ........A.T.CA.....GT-...TG.........A....G.....G..TC...T..-.....T..G..-...A--TAA
Apoly-Jap2 ........A.T.CA.....GT-...TG.........A....G.....G..TC...T..-.....T..G..-...A--TAA
Apoly-Mau1 ........A.T.CA....GGT-...TG.........A...AG.....G..TC...T..-.....T..G..-...A--CAA
Apoly-Mau2 ........A.T.CA....GGT-...TG.........A...AG.....G..TC...T..-.....T..G..-...A--CAA
Apoly-NZ1 ..........T.CA.....GT-...TG.........A..........G..TC...T..-........G..-....--TGA
Apoly-NZ2 ..........T.CA.....GT-...TG.........A..........G..TC...T..-........G..-....--TGA
Apoly-Phi1 ..........T.CA.....GT-...TG.........A...AG..G..G..TC...T..-........A..-...A--TAA
Apoly-Sey1 ..........T.CA.....GT-...TG.........A....G..G..G..TC...T..-........A..-...A--TA.
Apoly-Sey2 ..........T.CA.....GT-...TG.........A....G..G..G..TC...T..-........A..-...A--TA.
Arega-Pan1 .........C..........-.............A.......G.............-.G......G..-....----.
Ascop-Jap1 ..........T.CA.....GT-...TG..T......A...AG.....G..TC...TCG-........G..-...T--TAA
Aside-Pan1 ................G..-.....................................-.G.........-....----.
Asp1-Fij1 ..........T.CA.....GT-...TG..T......A...AG.....A..TC...TC.-........G..-...T--TAA
Asp2-Fij1 .A.....TA.T..A...G.TGT.......T.....GA......AG..A.......T.T-GG......G..-...A----A
Asp3-Fij1 .A......A.T..A..AGGTAT.C...........GA......AG.AA....T..T..--G...T..A..-...A-----
Asp3-Fij2 .A......A.T..A..AGGTAT.C...........GA......AG.AA....T..T..--G...T..A..-...A-----
Asp3-Fij3 .A......A.T..A..AGGTAT.C...........GA......AG.AA....T..T..--G...T..A..-...A-----
Asp4-Fij1 ........A.T..A..A..TAT...T.........GA..T...AG..A....C..T..-.G...T..A..-...------
Asp4-Fij2 ........A.TC.A..AG.TAT...T.........GA..T...AG..A....C..T..-.G...T..A..-...------
Asp5-Ton1 ..........T..A..A..TAT...T.........GAG.T...AG..A....C..T..-.G...T..G..-...------
Asp6-Phi1 ........A.T.CAA....-TT...TG.........A......AG..G..TC..TT..-GG...T..A..-...--AGGA
Asp7-Aus1 ..G...C..GT.-A..CG.AAG...TG.............A...G..A...C...T.G-.....T.....-....AGGG.
Aspin-Cor1 .........CT.CA......-G.............GA...A..AG.GG.......T..-.G......G..G.G..----.
Aspin-Cre1 .........CT.CA......-G.............GA...A..AG.GG.......T..-.G......G..G.G..----.
Atris-Oah1 .....C..ACT.CA.....GT-...TG.........A...AG.....G..TC......-.....T..G..-...A--TAA
Avapp-Bru1 ........AGT.CA.....GT-...TG..............G..G..G..TC...T.--.....T..G..-....--TAA
Averr-Cal1 ..................-.....................................-..........-....----.
Averr-Cal2 ..................-.............A.......G.............-.G......G..-....----.
Averr-Cal3 ..................-.............A.......G.............-.G......G..-....----.
Averr-Pan1 ..........T......G.T-.....................G.C...........-.G......G..-....----.
Azebr-PNG1 ..G...C..GT.-A...G.AAG...TG.............A...G..A...C...T.G-.....T.....-....AGGG.
Astro-Naz1 .A..C...AG..GAAT...TG.G..G.CAA...T......ACG.GG.A.AT..TT...TGGC.....A..-.T.A----.
Astro-Ton1 ....C...AG..GAAT...TG.A..G..AA...T......ACGAGG.A.GT..TT...TGGC.....A..-.T.A----.
Cphor-Mar1 T.......AGT.CA......TT.ATT....AA.C...G..AGG.G..AA......T.TTGGC....T...-..AA-----
Cpleu-Fij1 ...A......T.CA.A..GAGT...T........AG..G.A.GAG..AGGT..T.TC.AGG.AT.CTATACT.CCGAAGT
Pseud-Pan1 .......TCG.AGA.T...TG.G..G......AT.TCG...CGA...GGGT..TT..-TGGC..T.T....-.T-A----A
Tpeni-SDi1 T.......AGT..A...G.TGT.AAG....AA...G.....GGA...AAG..G..T.TA.GC....GA..-..A.-----
Tsubi-Far1 ........CG.GAA.T...TG.G..G..AA...T.GT...ACGAG..G.GT..TT..-TGGC..T.T....-.T-.----A


[1111111111111111222222222222222222222222222222222222222222222222222223333333333]
[4444555555566799000001111122223333344444445566666667777777788888899990000001111]
[6789034567912959012560145925670134671234567561245789014567890567895689012346023 5]
Aoers-Pan1 TTTTCGGTTTATTTAAACTTAAAATACTTGTATGATGTTAGGTGTTTGGTAGGTTGTTTTCACCTTCTTGTTGGCTTTAC
Aafri-STo1 .......C......T.....C.....A.G.....CT.....
Aalli-Col1 .-.....A......G.....A.....
Aamer-Flo1 .-.....A......G.....A.....
Aanti-Col1 .....A.....CC.....G.TA.....A.....
Aanti-Pan1 .....A.....CC.....G.TA.....A.....
Aaran-Cre1 ...C.....G.....A.....TCA.....A.....
Aaran-Cre2 ...C.....A.....TCA.....AA.....
Aaran-Cre3 ...C.....A.....TCA.....A.....
Aaran-Far1 ...C.....A.....TCA.....A.....
Aaran-Far2 ...C.....A.....TCA.....A.....
Aaran-Far3 ...C.....A.....TCA.....A.....
Aaran-Kav1 ...C.....A.....TCA.....A.....
Aaran-Kav2 ...C.....A.....TCA.....A.....
Aaran-Kav3 ...C.....A.....TCA.....A.....
Aaran-Mad1 ...C.....TCA.....A.....
Aaran-Mad2 ...C.....A.....TCA.....A.....
Aaran-Mad3 ...C.....A.....TCA.....AA.....
Aaran-Sar1 ...C.....A.....TCA.....A.....
Aaran-Sar2 ...C.....A.....TCA.....A.....
Aaran-Sar3 ...C.....A.....TCA.....AA.....
Aarma-Mex1 .....??????????????????????????????????
Aarti-Pan1 .....A.....CC.....G..A.....
Aarti-Pan2 .....A.....CC.....G..A.....
Aarti-Pan3 .....A.....CC.....G..A.....
Aarti-SCa1 .C.....C.....A.....C.....G.....AA.....
Aarti-SCa2 .C.....C.....A.....C.....G.....AA.....
Abisp-Cor1 A.....C....C.TG.....A.....CA.....A.....G.
Abisp-Cre1 A.....C....C.TG.....AA.....CA.....A.....G.
Abisp-Cre2 A.....C....C.TG.....A.....CA.....A.....G.
Abisp-Cre3 A.....C....C.TG.....AA.....C.....A.....G.
Abisp-Her1 A.....C....C.TG.....AA.....CA.....A.....G.
Abisp-Her2 A.....C....C.TG.....AA.....CA.....A.....G.
Abisp-Her3 A.....C....C.TG.....AA.....CA.....A.....G.
Abisp-Sar1 A.....C....C.TG.....A.....CA.....A.....G.
Abisp-Sar2 A.....C....C.TG.....A.....CA.....A.....G.
Abisp-Sar3 A.....C....C.TG.....AA.....CA.....A.....G.
Acing-Pan1 .....T.....A.....A.....C.....
Acomp-Flo1 .G.....C.....A.G..T.....C.....A.....
Acomp-Flo2 .G.....C.....A.G..T.....C.....A.....
Acomp-Flo3 .G.....C.....A.G..T.....C.....A.....
Adupl-Bim1 .....C.....G.....AA.C....
Adupl-Bim2 .....C.....G.....AA.C....
Adupl-Flo1 .....C.....G.....AA.C....
Adupl-SCa1 .....C.....G.....AA.C....
Adupl-SCa2 .....C.....G.....AA.C....
Aerin-Pan1 .....G.....A.....C.....WA.....
Agran-Aus1 .....A.....T.....T.....A.....A.GG...TA...T.....T.....AC.A.....
Aindi-Bru1 .....T.GG.....A.....CT.....G.T..A...T.....T.-.....A.C....
Aindi-Pak1 .....T.GG.....A.....CT.....G.T..A...T.....-.-.....A.C..??
Aindi-Pak2 .....T.GG.....A.....CT.....G.T..A...T.....-.-.....A.C....
Aindi-Tha1 .....T.GG.....A.....CT.....A..G.T..A...T.....T.-.....A....??
Airre-Far1 G.....C....C.TGG.....A.....A.....A.G.....T.....A.....
Airre-Far2 G.....C....C.TGG.....A.....A.....A.G.....T.....A.....
Airre-Far3 G.....C....C.TGG.....A.....A.....A.G.....T.....A.....
Airre-Kav1 G.....C....C..G.....T..A.....G.....C.A..A.....
Airre-Kav2 G.....C....C..G.....T..A.....G.....C.....A..A.....
Airre-Kav3 G.....C....C..G.....T..A.....G.....C.....C.A..A.....
Airre-Mad1 G.....C....C..G.....T..A.....G.....C.....A..A.....
Airre-Nor1 ??
Airre-Sar1 G.....C....C..G.....T..A.....G.....C.....A..A.....
Airre-Sar2 G.....C....C..G.....T..A.....G.....C.A..A.....
Airre-Sar3 A.....C....C.TG.....T..A.....G.....C.....A.....
Ajava-Bru1 ...A..A.....T.GG.....A.....C.....A.....AA..TA.....TT.....A.....
Ajons-Sar1 -...T..C.C....T....A.G.....G.A.....C.....A.....A.....
Ajons-Sar2 -...T..C.C....T....A.G.....G.A.....C.....A.....A.....
Ajons-Sar3 -...T..C.C....T....A.G.....G.A.....C.....A.....A.....
Ajons-Sar4 -.....C.C....T....A.G.....G.A.....C.....A.....A.....
Ajons-Tso1 -...T..C.C....T....A.G.....G.A.....C.....A.....A.....
Ajons-Tso2 -...T..C.C....T....A.G.....G.A.....C.....A.....A.....
Ajons-Tso3 -...T..C.C....T....A.G.....G.A.....C.....A.....A.....
Alate-Jap1 .A....A.....T.....A.....A.AG....A..CT.....T.....A.....
Amarg-Col1 .....AA.....
Amarg-Pan1 .....G.....????????

```
Amarg-Pan2 ..........................................................................AA......
Amona-Aus1 ......A.............T.....T.........A......A.GG...TA...T.......T.-...AC.A...A???
Aniti-Col1 .G.................................C.....A.G..T.......................A......
Aplat-Cyp1 A......C....C.TG............AA......AC....CA..........CA.....................G.
Aplat-Cyp2 A......C....C.TG............AA......AC....CA..........CA.....................G.
Aplat-Cyp3 A......C....C.TG............AA......AC....CA..........CA.....................G.
Aplat-LaH1 A......C....C.TG............AA......AC....CA..........CA.....................G.
Aplat-Sar1 A......C....C.TG............AA......AC....CA..........CA.....................G.
Aplat-Sar2 A......C....C.TG............AA......AC....CA..........CA.....................G.
Aplat-Sar3 A......C....C.TG............AA......AC....CA..........CA.....................G.
Aplat-Tou1 A......C....C.TG............AA......AC....CA..........CA.....................G.
Apoly-Fij1 G.....A.........G...T.............A.........A.GG...T...CT..............C.AA......
Apoly-Jap1 G....A......C.......T..............G........A.GG.T.....CT..............AC.TT......
Apoly-Jap2 G....A......C.......T..............G........A.GG.T.....CT..............AC.TT......
Apoly-Mau1 G....A......C.......T..............G........A.GC.T.....CT..............AC.TT.C....
Apoly-Mau2 G....A......C.......T..............G........A.GC.T.....CT..............AC.TCTC????
Apoly-NZ1  A....A......C.......T.........................GG.......AT.....................
Apoly-NZ2  A....A......C.......T.........................GG.......AT.....................
Apoly-Phi1 G....A......C.......T.....................A.GG.......CT........C.....C.A???????
Apoly-Sey1 G....A......C.......T.....................A.GG.......CT........C.....C.A.......
Apoly-Sey2 G....A......C.......T.....................A.GG........T........C.....C.A.......
Arega-Pan1 ..C.......................T.............A...........A.............A.......
Ascop-Jap1 G....A........T.G...T.....T..A......A......A.AGA...A..CT.......T.......C........
Aside-Pan1 ..........................................A.....................................
Asp1-Fij1  A....A...C....TG....T.....T.........A......A.CG....A..CT.......T.......C........
Asp2-Fij1  A..........C..............................T.CGT..TA...T..G....T.....A..........
Asp3-Fij1  G...........C......C.........A......A..G.....GAAA......T.AA.....................
Asp3-Fij2  G...........C......C.........A......A........GAAA......T.AA.....................
Asp3-Fij3  G...........C......C.........A......A..G.....GAAA......T.AA.....................
Asp4-Fij1  G.............................AA..............AAAA......T..A..........T..........
Asp4-Fij2  G.............................AA..............AAAA......T..A..........T..........
Asp5-Ton1  G...T........C.....C.........AA......A........GAAAC.....T..A.............A.......
Asp6-Phi1  G.....A.............T.............A..A......A.GG...T...CT................A...????
Asp7-Aus1  ......A.............T.GG...........C.........G.....AA..TA.....TT.........AA......
Aspin-Cor1 A......C.C..C.TG..........T..A..............A.G.......C...C..T.........CCA.......
Aspin-Cre1 A......C.C..C.TG..........T..A..............A.G.......C...C..T.........CCA.......
Atris-Oah1 AC...A......CC......T.............G........A.GG.T.....CT..............G.TCT-????
Avapp-Bru1 G....A..A...........T...C.........G...C......GG...G...CTA...........G--.AA...G..
Averr-Cal1 ............S..............AT...........A..........??????????????????????????
Averr-Cal2 ....................................G...A.....................................
Averr-Cal3 ....................................G...A.....................................
Averr-Pan1 ............C..............................A..........C.................A.......
Azebr-PNG1 ......A.............T.GG...........C.........G.....AA..TA.....TT.........AA......
Astro-Naz1 .C.CTA....GC..G.GGCAC.GG..T..A.....C....A..A.G..T..A.CG.GAGC.G.TAA...A..-.......
Astro-Ton1 .C.CTA....GC..G.GGCAC.GG..T..A.....C....A..A.G..T..A.CG.GAGC.G.TAG...A..T.......
Cphor-Mar1 ..C.........C.TG....C.......A..........TA..A.G...........AGCT...AA.......A......
Cpleu-Fij1 ..............T.....C.....T..A...........A.A.G..T..AA....AA....T........AA.....T
Pseud-Pan1 ...CTAA...GC....GGCAC.....T..A..............G..T..AACCAAAGC.GTTGG..GAG.TA......
Tpeni-SDi1 ......A.......TG...A........GA......A......A.AGT......CA..G.....G.......A.......
Tsubi-Far1 .C.CTAA...GC....GACAC.G...T..A.............ACG.AT..AAC.AAAGC.GTTA....AA.T..C....
```



```
Amarg-Pan2 ........GGG.-.................................A.............T..--....CT..TC.T.G
Amona-Aus1 ??????????????????????????.-............C.............A.....T.CA--...AA..TA.T.G
Aniti-Col1 .....A..GGG.-....A....-...??????????????????????????A..G..TT..--..........T..
Aplat-Cyp1 .C.....AGGGG-...T.T.........................A.........A.....T..--.....T...C...G
Aplat-Cyp2 .C.....AGGGG-...T.T.........................A.........A.....T..--.....T...C...G
Aplat-Cyp3 .C.....AGGGG-...T.T.........................A.........A.....T..--.....T...C...G
Aplat-LaH1 .C......GGGG-...T.T.........................A.........A.....T..--.....T...C...G
Aplat-Sar1 .C.....AGGGG-...T.T.........................A.........A.....T..--.....T...C...G
Aplat-Sar2 .C.....AGGGG-...T.T.........................A.........A.....T..--.....T...C...G
Aplat-Sar3 .C.....AGGGG-...T.T.........................A.........A.....T..--.....T...C...G
Aplat-Tou1 .C.....AGGGG-...T.T.........................A.........A.....T..--.....T...C...G
Apoly-Fij1 .G.....TAG..A....AT.......??????????????????????????????????????????????????
Apoly-Jap1 .G.....CAG..-...GAG...-...?????????????????.....A..A..TC.--.....T..A.....
Apoly-Jap2 .G.....CAG..-...GAG...-...?????????????????....CG.....A..A..TC.--.....T..A.....
Apoly-Mau1 .G.....CAG..-...GAG.....G.?.............................A..A..TC---.....T..AC...G
Apoly-Mau2 ????????????????????????????????????.....A..A..TC---.....T..AC...G
Apoly-NZ1  .G....GCAG..-.T..AG...-.............T.............C..A.....T.---.....T.CAC...G
Apoly-NZ2  .G....GCAG..-.T..AG...-.............T.............A.....T.---.....T.CAC...G
Apoly-Phi1 ????????????????????????????????????.....A..A..T.---.....T..AC...G
Apoly-Sey1 .G.C...TA...-....AG...-...???????????.....A.....T.---.....T..AC.T.G
Apoly-Sey2 .G.C...TA...-....AG.......???????????.....A....TT.---.....T..AC.T.G
Arega-Pan1 .......T....-T...........???????????.....A..T..--.........C...G
Ascop-Jap1 .A.....TA..--.T..AT...-..G.............................A.....T.---.....T.CAC.T.G
Aside-Pan1 ...........-.....................................................--.............
Asp1-Fij1  .A.....CA...-.T..AT.......??????????????????????????????????????????????????
Asp2-Fij1  ....G..AA...-...........???????????.....C...A.....T..--....G....C.T..
Asp3-Fij1  A.......GG.G-....?????????????????????T.............G.A.....T..--.....T..AC...G
Asp3-Fij2  A.......GG.G-...........????????????T.............G.A.....T..--.....T..AC...G
Asp3-Fij3  A.......GG.G-...........????????????????????????????????????????????????????
Asp4-Fij1  ........AG.G-....A........??????????????????????????????????????..TC..CGTGG
Asp4-Fij2  ........AG.G-....A........??????????????????????????????????????..TC..CGTGG
Asp5-Ton1  ........GG.G-..A.A........??????????????????????????????????????????????????
Asp6-Phi1  ??????????????????????????.....................A...A..C.A-..G..T...C.T.G
Asp7-Aus1  .G....GTA...-T...AG.......??????????????????.....G.....A....TT.---......AC.A.T..
Aspin-Cor1 .A......AGG.-...TAT...-.................C..A.............T..--........CT.TT..
Aspin-Cre1 .A......AGG.-...TAT...-.................C..A.............T..--........CT.TT..
Atris-Oah1 ????????????????????????????????????.....A..A..T.A--.....T..AC...G
Avapp-Bru1 .G....GCA...-.T..AG...-.??????????????????.....G.....A.....TC---.....T.-AC.T.G
Averr-Cal1 ????????????????????????T.....................................--.............
Averr-Cal2 .......??????????????????.....................................T..--........A.....
Averr-Cal3 .......??????????????????????????.....................T..--........A.....
Averr-Pan1 ........-G...-.....................................................T..--.........C....
Azebr-PNG1 .G....GTA...-....AG.......???????????.....A....TT.---......AC.A.T..
Astro-Naz1 A.....TTA...A?????????????????????????...C.CT.A.T.......AT.T...CCAAGAGC..A.T.GA.G
Astro-Ton1 A.C...TTA...AGA....GA.....??????????????????????????????????????????????????
Cphor-Mar1 A......AGT..-..AG.T.......??????????????????????????????????????????????????
Cpleu-Fij1 ......GTGG..-T...AT.......??????CTTGG.....C.......C....A..A..T..T-...T.T.....T.G
Pseud-Pan1 ......TAA...TAA..A.GA.TC......T........AC.C.TAA........AT.T..T..AT...T.T..T.GA.G
Tpeni-SDi1 A..-..?????????????????????????????.....AT.A..AAATA...T.T..TC.T.G
Tsubi-Far1 ......TAA...TAA....GA.........T.........C.C..AA........AT.T.....AT..G-CT..T.GA.G
```

```
[          5555555555555555666666666666666666666666666666666666666666666666666666666666666666666666]
[          6666667777888889000001111111111222222222233333333334444444445555555555666666666667]
[          1234574678012890567890123456789012345678901234567890123456780123456789012345678 0]
Aoers-Pan1 AAAT-GGTTGTTTTGGACAACTAA-TTAGCAGTA-AACTTTTAATT-CCGAGTAAAA-TCAAATTAATAACTCTAATTCA
Aafri-STo1 ....-.AG.....C.....TT...-....A.A.T-.GT.....TCC-TTA.......-CT.................CA.
Aalli-Col1 ....-.T..............C.-...A.....-..G....GCC.-T.A.......-C.G.....G.........CC..
Aamer-Flo1 ....-.T..............C.-...A.....-..G....GCC.-T.A.......-C.G.....G.........CC..
Aanti-Col1 ..G.-A.............A.T.-........T-G.....CGCC.-TTA.A.....-....-........C....CC..
Aanti-Pan1 ..G.-A.............A.T.-........T-G.....CGCC.-TTA.A.....-....-........C....CC..
Aaran-Cre1 ..G.--.G.A...C.....TTCT.-A...T...T-..TC....T.C-.TAG......T.AG.....TC.......T.AT.
Aaran-Cre2 ..G.--.G.A...C.....TTCT.-A...T...T-..TC....T.C-.TAG......T.AG.....TC.......T.AT.
Aaran-Cre3 ..G.--.G.A...C.....TTCT.-A...T...T-..TC....T.C-.TAG......T.AG.....TC.......T.AT.
Aaran-Far1 ..G.--.G.A...C.....TTCT.-A...T...T-..TC....T.C-.TAG......T.AG.....TC.......T.AT.
Aaran-Far2 ..G.--.G.A...C.....TTCT.-A...T...T-..TC....T.C-.TAG......T.AG.....TC.......T.AT.
Aaran-Far3 ..G.--.G.A...C.....TTCT.-A...T...T-..TC....T.C-.TAG......T.AG.....TC.......T.AT.
Aaran-Kav1 ..G.--.G.A...C.....TTCT.-A...T...T-..TC....T.C-.TAG......T.AG.....TC.......T.AT.
Aaran-Kav2 ..G.--.G.A...C.....TTCT.-A...T...T-..TC....T.C-.TAG......T.AG.....TC.......T.AT.
Aaran-Kav3 ..G.--.G.A...C.....TTCT.-A...T...T-..TC....T.C-.TAG......T.AG.....TC...C...T.AT.
Aaran-Mad1 ..G.--.GAA...C.....TTCT.-A...T...T-..TC....T.C-.TAG......T.AG.....TC.......T.AT.
Aaran-Mad2 ..G.--.GAA...C.....TTCT.-A...T...T-..TC....T.C-.TAG......T.AG.....TC.......T.AT.
Aaran-Mad3 ..G.--.GAA...C.....TTCT.-A...T...T-..TC....T.C-.TAG......T.AG.....TC.......T.AT.
Aaran-Sar1 ..G.--.G.A...C.....TTCT.-A...T...T-..TC....T.C-.TAG......T.AG.....TC.......T.AT.
Aaran-Sar2 ..G.--.GAA...C.....TTCT.-A...T...T-..TC....T.C-.TAG......T.AG.....TC.......T.AT.
Aaran-Sar3 ..G.--.G.A...C.....TTCT.-A...T...T-..TC....T.C-.TAG......T.AG.....TC.......T.AT.
Aarma-Mex1 ....-.................-.........-....C......-.........-...................C..
Aarti-Pan1 ....-A.............A.T.-........T-......CGCC.-TGA.......-....-........C....CC..
Aarti-Pan2 ....-A.............A.T.-........T-......CGCC.-TGA.......-....-........C....CC..
Aarti-Pan3 ....-A.............A.T.-........T-......CGCC.-TGA.......-....-........C....CC..
Aarti-SCa1 ....-A..........G..GA.CG-.........-...C....CC.-TTA.......-....-........C....CC..
Aarti-SCa2 ....-A..........G..GA.CG-.........-...C....CC.-TTA.......-....-........C....CC..
Abisp-Cor1 .C..-.AGG....C....G.T.TG--..A....T-...C..C.CC.-TT........-C.G..............CCA.
Abisp-Cre1 .C..-.AAG....C....G.T.TG--..A....T-...C..C.CC.-TT........-..G..............CCA.
Abisp-Cre2 .C..-.AGG....C....G.T.TG--..A....T-...C..C.CC.-TT........-C.G..............CCA.
Abisp-Cre3 .C..-.AAG....C....G.T.TG--..A....T-...C..C.CC.-TT........-..G..............CCA.
Abisp-Her1 .C..-.AAG....C....G.T.TG--..A....T-...C..C.CC.-TT........-..G..............CCA.
Abisp-Her2 .C..-.AAG....C....G.T.TG--..A....T-...C..C.CC.-TT........-..G..............CCA.
Abisp-Her3 .C..-.AAG....C....G.T.TG--..A....T-...C..C.CC.-TT........-..G..............CCA.
Abisp-Sar1 .C..-.AAG....C....G.T.TG--..A....T-...C..C.CC.-TT........-C.G..............CCA.
Abisp-Sar2 .C..-.AAG....C....G.T.TG--..A....T-...C..C.CC.-TT........-C.G..............CCA.
Abisp-Sar3 .C..-.AAG....C....G.T.TG--..A....T-...C..C.CC.-TT........-..G..............CCA.
Acing-Pan1 ..G.-.....C..........T.-.........-..T.....CCC-T-A.......-C.......GA........CC..
Acomp-Flo1 ..G.-........C........C.-.........-.......GCC.-..A.......-............C....CC..
Acomp-Flo2 ..G.-........C........C.-.........-.......GCC.-..A.......-............C....CC..
Acomp-Flo3 ..G.-........C........C.-.........-.......GCC.-..A.......-............C....CC..
Adupl-Bim1 ..G.-A.............A.TG-...A..A.T-..T...CGCC.-TTA.......-.T..-........C....CC..
Adupl-Bim2 ..G.-A.............A.TG-...A..A.T-..T...CGCC.-TTA.......-.T..-........C....CC..
Adupl-Flo1 ..G.-A.............A.TG-...A..A.T-..T...CGCC.-TTA.......-.T..-......G.C....CC..
Adupl-SCa1 ..G.-A.............A.TG-...A..A.T-T.T...CGCC.-TTA.......-.T..-........C....CC..
Adupl-SCa2 ..G.-A.............A.TG-...A..A.T-..T...CGCC.-TTA.......-.T..-........C....CC..
Aerin-Pan1 ..G.-A...........G.....-.......A.-...C..C.C.C-..........-C.........G.......CC..
Agran-Aus1 ..----A.AT...C.AGT.C....-...A.CA.C-T.T...C--..A.TA.......-CT......-.....T.T...AT
Aindi-Bru1 ..---.AAAT......G.G.T.T.-C..A..A.T-C.......-CCGTTT.......-.T.T....-A....T....AA.
Aindi-Pak1 ..---.AGAT......G...T.T.-C..A..A.C-T.T.....-CCG.TT.......-.T.T....-A....T....AA.
Aindi-Pak2 ..---.AGAT......G...T.T.-C..A..A.C-T.T.....-CCG.TT.......-.T.T....-A....T....AA.
Aindi-Tha1 ..---..G.T......G.G.T.T.-A..A..A.C-TGT.....-CCATTTG......-.T.T....-A....T....AA.
Airre-Far1 ....-.AAAA...C..G..GT.CG--.......T-.......GT..-TTAG......-CTG.......G.....T.C.AG
Airre-Far2 ....-.AAAA...C..G..GT.CG--.......T-.......GT..-TTAG......-CTG.......G.....T.C.AG
Airre-Far3 ....-.AAAA...C..G..GT.CG--.......T-.......GT..-TTAG......-CTG.......G.....T.C.AG
Airre-Kav1 ....-.AAAA...C..G...T.TG--.......T-..T.C...C..-TTA.......-CTG.....T.......T.C.AG
Airre-Kav2 ....-.AAAA...C..G...T.TG--.......T-..T.C...C..-TTA.......-CTG.....T.......T.C.AG
Airre-Kav3 ....-.AAAA...C..G...T.TG--.......T-..T.C...C..-TTA.......-CTG.....T.......T.C.AG
Airre-Mad1 ....-.AAAA...C..G...T.TG--.......T-..T.C...C..-TTA.......-CTG.....T.......T.C.AG
Airre-Nor1 ....-.AGAA...C..G...T.TG--.......T-..T.....C..-TTA.......-CT......T.......T.C.AG
Airre-Sar1 ....-.AAAA...C..G...T.TG--.......T-..T.C...C..-TTA.......-CTG.....T.......T.C.AG
Airre-Sar2 ....-.AAAA...C..G...T.TG--.......T-..T.C...C..-TTA.......-CTG.....T.......T.C.AG
Airre-Sar3 ....-.AGAA...C..G...T.TG--.......T-..T.....C..-TTA.......-CT......T.......T.C.AG
Ajava-Bru1 ..---..GAT...C..G..TT.T.G...A..ACTCTGT....GT.-GTT......G.-.TTT.....-..G.CT...CCA.
Ajons-Sar1 ..-.A.A.ATA..C........C.-........T-..T.C.C.T..-TTA.......-CA.......C...C......A.
Ajons-Sar2 ..-.A.A.ATA..C........C.-........T-..T.C.C.T..-TTA.......-CA.......C...C......A.
Ajons-Sar3 ..-.A.A.ATA..C........C.-........T-..T.C.C.T..-TTA.......-CA.......C...C......A.
Ajons-Sar4 .-G.A.A.ATA..C........C.-........T-..T.C.C.T..-TTA.......-CA.......C...C......A.
Ajons-Tso1 .-G.A.A.ATA..C........C.-........T-..T.C.C.T..-TTA.......-CA.......C...C......A.
Ajons-Tso2 .-G.A.A.ATA..C........C.-........T-..T.C.C.T..-TTA.......-CA.......C...C......A.
Ajons-Tso3 .-G.A.A.ATA..C........C.-........T-..T.C.C.T..-TTA.......-CA.......C...C......A.
Alate-Jap1 ..---.AAAT.C.CA....TTCT.-...A..A.T-T....C..C..A.TA.......-CT.T....-...TCT.....T.
Amarg-Col1 ..G.-.................C.-.........-..A..C.GTAC-.-A.......-CT................CGT.
Amarg-Pan1 ..G.-.................C.-.........-..A..C.GTAC-.-A.......-CT................CGT.
```

```
Amarg-Pan2 ..G.-.................C.-.........-..A..C.GTAC-.-A.......-CT................CGT.
Amona-Aus1 ..----A.AT...C.AGT.C....-...A.CA.C-T.T...C--..A.TA.......-CT......-.....T.T...AT
Aniti-Col1 ..G.-.......-C........C.-.........-......-GCC.-..........-.............C....CC..
Aplat-Cyp1 ....-.AAAA...C....G.T.TG--.......T-.......GCCC-TTA.......-..G...............C.A.
Aplat-Cyp2 ....-.AAAA...C....G.T.TG--.......T-.......GCCC-TTA.......-..G...............C.A.
Aplat-Cyp3 ....-.AAAA...C....G.T.TG--.......T-.......GCCC-TTA.......-..G...............C.A.
Aplat-LaH1 ....-.AAAA...C....G.T.TG--.......T-.......GCCC-TTA.......-..G...............C.A.
Aplat-Sar1 ....-.AAAA...C....G.T.TG--.......T-.......GCCC-TTA.......-..G............C..C.A.
Aplat-Sar2 ....-.AAAA...C....G.T.TG--.......T-.......GCCC-TTA.......-..G............C..C.A.
Aplat-Sar3 ....-.AAAA...C....G.T.TG--.......T-.......GCCC-TTA.......-..G............C..C.A.
Aplat-Tou1 ....-.AAAA...C....G.T.TG--.......T-.......GCCC-TTA.......-..G............C..C.A.
Apoly-Fij1 ??????????????????????????????????????????????????????????????????????????????
Apoly-Jap1 .G---..AAA......G.GC..T.-C..A..A.T-T.T.....T.A-.TAG.......-.T.T....-....CT...C.TG
Apoly-Jap2 .G---..AAA......G.GC..T.-...A..A.T-T.T.....T.A-.TAG.......-.T.T....-....CT...C.TG
Apoly-Mau1 .G---..GAT...C..G.GC..T.-...A..A.T-T.T.....T.A-.TAG.......-.T.T....-.....T...C..G
Apoly-Mau2 .G---..GAT...C..G.GC..T.-...A....T-T.T.....T.A-.TAG.......-.T.T....-.....T...C..G
Apoly-NZ1  ..---..GAT.........TT.CG-C..A..A.T-T-T.....T.-AT.A.......-...T....-.....T...C.T.
Apoly-NZ2  ..---A.GAT.........TT.CG-C..A..A.T-T-T.....T.-AT.A.......-...T....-.....T...C.T.
Apoly-Phi1 .G---..AAT...C..G..C..TG-...AA.A.T-T.T.....T..-.TAG.......-.T.T....-.....T...C.T.
Apoly-Sey1 .G---..AAT...C..G.GC..T.-...AA.A.T-TGT.....T..-.TAG.......-.T.T....-.....T...C.T.
Apoly-Sey2 .G---..AAT...C..G.GC..T.-...AA.A.T-TGT.....T..-.TAG.......-.T.T....-.....T...C.T.
Arega-Pan1 ....-A...T.............-.........-...CCCC.C.C-..........-..................C.A.
Ascop-Jap1 ..---CTGAT.........CT.T.-...A....T-T.T.C...T.-ATTA.......-...T....-.....T...C.A.
Aside-Pan1 ....-.................-.........-..........-..........-..................C..
Asp1-Fij1  ??????????????????????????????????????????????????????????????????????????????
Asp2-Fij1  ....--.GAA......G..C..C.-.C.AA-TA.T..T.C...T.C-TTA.A..G..-C.......-...TCT.G.C.TG
Asp3-Fij1  ....-A.AAC...C..G..TGCC.-C..A.TT.TTG.TAC.C.T..-.TAG.......-CT.G....-C..TCT.....TG
Asp3-Fij2  ....-A.AAC...C..G..TGCC.-...A.TT.TTG.TAC.C.C..-.TAG.......-CT.G....-C..TCT.....TG
Asp3-Fij3  ??????????????????????????????????????????????????????????????????????????????
Asp4-Fij1  ..-.--AAAC...C..G..T.AC.-C..A.CT.TTG.TA....TCC-..AG.......-CT.......-C..T.T....G..
Asp4-Fij2  ..-.--AAAC...C..G..T.AC.-C..A.CT.TTG.TA....TCC-..AG.......-CT.......-C..T.T....G..
Asp5-Ton1  ??????????????????????????????????????????????????????????????????????????????
Asp6-Phi1  ..-----.AAT......G..CACT.-...A..AAT-T.T..C..-.CAT.A........-CTTT....-...T.T....CGT
Asp7-Aus1  ..---..GAT...C..G..TT.T.G...A..ACT-TGT...-.T.-GTT......G.-.T.T....-...T.T...CCA.
Aspin-Cor1 ....-.AAAA...C.....GT.CG--....G..T-..T.....C..-TT.........-C.T.....T.......T.C...
Aspin-Cre1 ....-.AAAA...C..G..GT.CG--....G..T-..T.....C..-TT.........-C.T.....T.......T.C...
Atris-Oah1 .G---..AAT......G.CT..T.-...A..A.T-T.T..C..T.C-.T.G.......-.T.T....-.....T...C.TG
Avapp-Bru1 ..---..AAT...C..G..TT.TG-CC.A..ACT-T.T.C.C.T.-G..A........-.T.T....-C....T...C.T.
Averr-Cal1 ....-.................-.........-..........-..........-.T.....................
Averr-Cal2 ....-.................-.........-......C.G..-..........-..........G.......CC..
Averr-Cal3 ....-.................-.........-......C.G..-..........-..........G.......CC..
Averr-Pan1 ..C.-A.............G.-...A...A.-.......C.C.C-..........-C.................CT.
Azebr-PNG1 ..---..GGT...C..G..TT.T.A...A..ACT-TGT.....T.-GTT.........-.T.T....-..GT.T...C.A.
Astro-Naz1 TTC.T.TGAAA..C.......AC.TA..TT.T.TT..A..A..C..ATTA--CT...AAT...AACCA.TACT..CAC.T
Astro-Ton1 ??????????????????????????????????????????????????????????????????????????????
Cphor-Mar1 ??????????????????????????????????????????????????????????????????????????????
Cpleu-Fij1 ..TCTAA..T...C...TGTT...CA.TAT.C.TA...A..A.C..TTTA.A.....AAT..TC.TTA..ACA.TGC.A.
Pseud-Pan1 .TC.TATAAAA.AC.....TAA..CAC.C..TCTT..G..A..-..TTT.TA.T..GAAT...AA....TA.T..TA..T
Tpeni-SDi1 ..TCTAA.AAA..C.AG...T...AC.TTTTA.TTT.AA....T..CTTT.A.....--......G.C.TA.......TT
Tsubi-Far1 .TC.TATAAT...C......AA..CAC.T..C.TT..A..C..C..TTTATT.TC.GAAT...AA....TA.T..TC.TT
```

```
[          6666666777777777777777777777777777777777777777777777888888888888888888888888899]
[          7777788000001111111222333333444444555555666677777788911133555666666677788899990 0]
[          1234635237890345689789023458356789012346345801238928023489039124568914703602581 4]
Aoers-Pan1 TAGCTGCTACTTTCCCCCCGAGGTGATGAATTCTAAAAGGTTACTTCCTGAACTGAAAGTCCCACGAAAACCAAGCATTT
Aafri-STo1 ....G.T...C...T......A.....A............C...CC.................A...T...T....A....
Aalli-Col1 ..........C...TT.....A...G.A...C.......................??????????????????????
Aamer-Flo1 ..........C...TT.....A...G.A...C.........................????????????????????
Aanti-Col1 ..........C...TT...A.A.....A...C.......................????T.T..........G.T.A..
Aanti-Pan1 ..........C...TT...A.A.....A...C.....................G...T.T............T.A..
Aaran-Cre1 A.....T.......TT.....A...GCA............C...C........???????T.......C......T.A..
Aaran-Cre2 A.............TT.....A...GCA............C...C........???????T.......C......T.A..
Aaran-Cre3 A.....T.......TT.....A...GCA............C...C........???????T.......C......T.A..
Aaran-Far1 A.....T.......TT.....A...GCA............C...C............????T.......C......T.A..
Aaran-Far2 A.....T.......TT.....A...GCA............C...C............????T.......C......T.A..
Aaran-Far3 A.....T.......TT.....A...GCA............C...C............????T.......C......T.A..
Aaran-Kav1 A.....T.......TT.....A...GCA............C...C............G...T.......C......T.A..
Aaran-Kav2 A.....T.......TT.....A...GCA............C...C............G...T.......C......T.A..
Aaran-Kav3 A.....T.......TT.....A...GCA............C...C............G...??????????????????
Aaran-Mad1 A.....T.......TT.....A...GCA............C...C............???????????????????????
Aaran-Mad2 A.....T.......TT.....A...GCA............C...C............???????????????????????
Aaran-Mad3 A.............TT.....A...GCA............C...C............????T.......C......T.A..
Aaran-Sar1 A.....T.......TT.....A...GCA............C...C............????T.......C......T.A..
Aaran-Sar2 A.....T.......TT.....A...GCA............C...C........???????T.......C......T.A..
Aaran-Sar3 A.............TT.....A...GCA............C...C............????T.......C......T.A..
Aarma-Mex1 ..............TT.....A.....A.........................???????????????????????
Aarti-Pan1 ..........C...TT...A.A.....A...C.......................????T.T............T.A..
Aarti-Pan2 ..........C...TT...A.A.....A...C.......................????T.T............T.A..
Aarti-Pan3 ..........C...TT...A.A.....A...C.....................G...T.T............T.A..
Aarti-SCa1 .........T....TT...A.A.....A.G.C.....................G...??????????????????????
Aarti-SCa2 .........T....TT...A.A.....A.G.C.....................G...??????????????????????
Abisp-Cor1 C...A.T...AC..T......A.....A..CA.....G..CC..CC.........??????????????????????????
Abisp-Cre1 C...A.T...AC..T......A.....A..CA.....G..CC..CC.......???????????????????????????
Abisp-Cre2 C...A.T...AC..T......A.....A..CA.....G..CC..CC.......???????????????????????????
Abisp-Cre3 C...A.T...AC..T......A.....A..CA.....G..CC..CC.......???????..A...T.T.......GC.C
Abisp-Her1 C...A.T...AC..T......A.....A..CA.....G..CC..CC..............??????????????????????
Abisp-Her2 C...A.T...AC..T......A.....A..CA.....G..CC..CC..............??????????????????????
Abisp-Her3 C...A.T...AC..T......A.....A..CA.....G..CC..CC..............??????????????????????
Abisp-Sar1 C...A.T...AC..T......A.....A..CA.....G..CC..CC..........????..A...T.T.......GC.C
Abisp-Sar2 C...A.T...AC..T......A.....A..CA.....G..CC..CC.......???????..A...T.T.......GC.C
Abisp-Sar3 C...A.T...AC..T......A.....A..CA.....G..CC..CC..........????..A...T.T.......GC.C
Acing-Pan1 ..........C...TT...A.A.....AG.CC..................G.....????.........T..G..T..CC
Acomp-Flo1 ..........C...TT.....A...G.A..CC........C.....................??????????????????????
Acomp-Flo2 ..........C...TT.....A...G.A..CC........C.....................??????????????????????
Acomp-Flo3 ..........C...TT.....A...G.A..CC........C.....................??????????????????????
Adupl-Bim1 ..............TT...A.A.....A...C.C.......C..............????T.T............T.A..
Adupl-Bim2 ..............TT...A.A.....A...C.........C..............???????T.T............T.A..
Adupl-Flo1 ..............TT...A.A.....A...C.........C..............G...??????????????????????
Adupl-SCa1 ..............TT...A.A.....A...C.........C..............G...??????????????????????
Adupl-SCa2 ..............TT...A.A.....A...C.........C..............G...??????????????????????
Aerin-Pan1 .........T....TT.....A.....A...C.........................????T..........T........
Agran-Aus1 ..A.A.T....C..TT.TTA.A.CA..A..AC........C...........T.......???????...TT.....A..
Aindi-Bru1 ..A.A.T....G..TT.TTA.A..A..A..AC........C...CC......T...????A..........T...T..C.
Aindi-Pak1 ..A...T...CG..TT.TTA.A..A..A..AC........C...CC......T.......A..........T...T..C.
Aindi-Pak2 ..A...T...CG..TT.TTA.A..A..A..AC........C...CC......T.......A..........T...T..C.
Aindi-Tha1 ..A...T....A..TT.TTA.A..A..A..AC............CC......T.......A..........T...T.CC.
Airre-Far1 ....A.T.......T......A...G.A..C.........CC..CC.......???????A.....C....T.T...A..
Airre-Far2 ....A.T.......T......A...G.A..C.........CC..CC.......???????A.....C....T.T...A..
Airre-Far3 ....A.T.......T......A...G.A..C.........CC..CC.......???????A.....C....T.T...A..
Airre-Kav1 ....A.....C...T......A.....A............C...CC..............TT..T.C...TT.G...A..
Airre-Kav2 ....A.....C...T......A.....A............C...CC..............??..T.C...TT.G...A..
Airre-Kav3 ....A.....C...T......A.....A............C...CC................A.T.T....T...T.A..
Airre-Mad1 ....A.....C...T......A.....A............C...CC..............????T.C...TT.G...A..
Airre-Nor1 ....A.........T......A.....A..C.........C...CC..............??????????????????????
Airre-Sar1 ....A.....C...T......A.....A............C...CC.......???????????????????????????
Airre-Sar2 ....A.....C...T......A.....A............C...CC.......???????T...T.T...TT.G...A..
Airre-Sar3 ....A.........T......A.....A..C.........C...CC.......???????T.....T...TT.G...A..
Ajava-Bru1 ..A.A.T...AAC.TT.TTA.A...A..AC........C...CC.GGC..TA..????A....................
Ajons-Sar1 ..ATA.T..T....TTT..A.A.....A............C...CC.......???????..A.T.T....T...T.A..
Ajons-Sar2 ..ATA.T..T....TTT..A.A.....A............C...CC.......???????..A.T.T....T...T.A..
Ajons-Sar3 ..ATA.T..T....TTT..A.A.....A............C...CC.......???????..A.T.T....T...T.A..
Ajons-Sar4 ..ATA.T..T....TTT..A.A.....A............C...CC.......???????..A.T.T....T...T.A..
Ajons-Tso1 ..ATA.T..T....TTT..A.A.....A............C...CC..............??????????????????????
Ajons-Tso2 ..ATA.T..T....TTT..A.A.....A............C...CC................A.T.T....T...T.A..
Ajons-Tso3 ..ATA.T..T....TTT..A.A.....A............C...CC................A.T.T....T...T.A..
Alate-Jap1 C...A.T.....G.TT.TTA.A..A..A..ACT..G....C...........T...????A.....T...T..C......
Amarg-Col1 ..............T......A.....A..C.........C..........C.???????????...T...T...T...C
Amarg-Pan1 ..............T......A...G.A..C.........C..........T.???????...........T...T...C
```

```
Amarg-Pan2 ..............T......A.....A..C.........C..........T....??????????????????????
Amona-Aus1 ..A.A.T....C..TT.TTA.A.CA..A..AC........C...........T.......A.T.........G.......
Aniti-Col1 ..........C...TT.....A...G.A..CC........CC.........--???????A............C..GC..
Aplat-Cyp1 ....A.T...AC..T......A...G.A............CC..CC..........????????????????????????
Aplat-Cyp2 ....A.T...AC..T......A...G.A............CC..CC........??????..A...T.C......TGA.C
Aplat-Cyp3 ....A.T...AC..T......A...G.A............CC..CC..........????..A...T.C......TGA.C
Aplat-LaH1 ....A.T...AC..T......A...G.A............CC..CC.......????????????????????????????
Aplat-Sar1 ....A.T...AC..T......A...G.A............CC..CC.......???????..A...T.C......TGA.C
Aplat-Sar2 ....A.T...AC..T......A...G.A............CC..CC.......???????..A...T.C......TGA.C
Aplat-Sar3 ....A.T...AC..T......A...G.A............CC..CC.......???????..A...T.C......TGA.C
Aplat-Tou1 ....A.T...AC..T......A...G.A............CC..CC.......????????????????????????????
Apoly-Fij1 ??????????????????????????????????????????????????????????????????????????????????
Apoly-Jap1 ..A.G.T........T.TT.....A..A.TCC........C.G.C.......T...????A.A............A.A..
Apoly-Jap2 ..A.G.T........T.TT.....A..A.TCC........C.G.C.......T...????A.A............A.A.C
Apoly-Mau1 ..A.A.T..T.C...T.TT.....A..A.T.C........C...C.......T.......??????????????????
Apoly-Mau2 ..A.A.T..T.C...T.TT.....A..A.T.C........C...C.......T........A.A.........G..A.A..
Apoly-NZ1  ....A.T..........TT.G...A..A.T..........C...C.......T???????A.A.......T....A.AC.
Apoly-NZ2  ....A.T..........TT.G...A..A.T..........C...C.......T???????A.A.......T....A.AC.
Apoly-Phi1 ....A.T..........TT.G...A..A.T.C........C...C.......T......?????????.......T.A.C
Apoly-Sey1 ....A.T..........TTAG...A..A.T.C........C...C.......T.....??A.A.........T...T.A.C
Apoly-Sey2 ....A.T..........TTAG...A..A.T.C........C...C.......T???????A.A.........T...T.A.C
Arega-Pan1 ..............TT.....A...G.AG..C.............................??????????.........
Ascop-Jap1 ....A.T...C...TTATTT.A..A..A.C.C........C...CC......TACG????A.........T....AGA..
Aside-Pan1 ..............TT.....A.....A..............................????.....................
Asp1-Fij1  ??????????????????????????????????????????????????????????????????????????????????
Asp2-Fij1  ..A.A.T.GTC...T......A.....A...CT.......C...................??????????????????....
Asp3-Fij1  .T..A.T........TT.....A.....A..C.........C...CC..............A.............A..C.C
Asp3-Fij2  .T..A.T........TT.....A.....A..C.........C...CC..............A.............A..C.C
Asp3-Fij3  ??????????????????????????????????????????????????????????A.............A..C.C
Asp4-Fij1  C.A.A.T.....C..TT..A.AA..G.A..C.........CC..................????????..........A.G
Asp4-Fij2  C.A.A.T.....C..TT..A.AA..G.A..C.........CC..................????????????????????
Asp5-Ton1  ??????????????????????????????????????????????????????????????????????.....A.A.A
Asp6-Phi1  A.A.A.T.......TTTTTA.A..A..A..AC........C...........TA..????A..........T.T...A..
Asp7-Aus1  ..A.A.T...AAC.TTTTTA.A..A..A..A.........C...CC......T...????A................AC.
Aspin-Cor1 ....A......C..T....AGA.....A..C.........C...CC.......???????......T.....GG...ACC
Aspin-Cre1 ....A......C..T....AGA.....A..C.........C...CC.......???????......T.....GG...ACC
Atris-Oah1 ..A.A.T........T.TTA....A..A.T.A........C...C.......T.......??????????????????
Avapp-Bru1 ..A.A.T......T...TTA..A.A..A.T..........C...C.......TA..????A.GT...........A....
Averr-Cal1 ..............T.....................................????....................
Averr-Cal2 ..............TT.....A.....A..CC.....................??????T?..................
Averr-Cal3 ..............TT.....A.....A..CC.........................T?..................
Averr-Pan1 .........T....T......A.....A.T.C.........................T.T................C
Azebr-PNG1 ..A.A.T...AAC.TTTTTA.A..A..A..A.........C...CC......T....G..???????...TT.....A..
Astro-Naz1 .T...ATC....A...T.TA...AAT.TGG..T..GG.AA.C.T.CT...TT........T.T..A..CT..C..A.ACG
Astro-Ton1 ??????????????????????????????????????????????????????????T.T.....TC..C..A.ACA
Cphor-Mar1 ??????????????????????????????????????????????????????????????????????????????????
Cpleu-Fij1 A.ATA.T..T..C.T....A.A...TAA...C......A.C..T............G.A.??????????????????
Pseud-Pan1 CT.T.ATC..........TA...AAT.TG..AT.G.G.A....T..T....T....G..GA.......TC..C..A...G
Tpeni-SDi1 ..A..AT..A..C.TTT.TA.A.CATC..............C.T..............T....A...TTTT.A.GA..
Tsubi-Far1 .T.T.ATC.T..C...T.TA...AAT.TG..GA.G...A....T..T....-..??????T...T...T...C.AA...G
```

```
[                                                  1111111111111111111111111111111111111111111111111]
[          9999999999999999999999999999999999900000000000000000000000000000000001111111111]
[          0011122222223334444455556667778899900000011122333344455666777788889999900000011112]
[          7803923456891470346923581470692514703569258140369258170362568146701369235891247O]
Aoers-Pan1 TCAATCGCAAAGCAAAATAACCATAACCTACGCACCACACCCTAGATTCATAAATCACCCTAAGACTAAACACACACTAC
Aafri-STo1 .T.TC...C...........TT.C.T........TT..T.TT..CT...TAT.TCA..A.A.....C...T.TC.T....
Aalli-Col1 ????????????????????????????????????????????????????????????????????????????????????
Aamer-Flo1 ????????????????????????????????????????????????????????????????????????????????????
Aanti-Col1 C.......C....T..CC...T.........AT..T....TTC.A.AC.G...G.A.T.........C.GG.......C..
Aanti-Pan1 C.......T...AC..CC...T.........AT..T....TTC.A.AC.G...G.A.T.........C.GG.......C..
Aaran-Cre1 ..C.A...C...TTT.T..G.....C..C.TT....G...TT..TTA.T.CT.....T....G...C.C.T..T.T...A
Aaran-Cre2 ..C.A...C...TTT.T..G.....C....TT...TG...TT..TTA.T.CT.....T....G...C.C.T..T.T...A
Aaran-Cre3 ..C.A...C...TCT.T..G.....C....TT...TG...TT..TTA.T.CT.....T....G...C.C.T..T.T...A
Aaran-Far1 ..C.A...C...TCT.T..G.....C....TT...TG...TT..TTA.T.CT.....T....G...C.C.T..T.T...A
Aaran-Far2 ..C.A...C...TCT.T..G.....C....TT...TG...TT..TTA.T.CT.....T....G...C.C.T..T.T...A
Aaran-Far3 ..C.A...C...TCT.T..G.....C....TT...TG...TT..TTA.T.CT.....T....G...C.C.T..T.T...A
Aaran-Kav1 ..C.A...C...TCT.T..G.....C....TT...TG...TT..TTA.T.CT.....T....G...C.C.T..T.T...A
Aaran-Kav2 ..C.A...C...TTT.T..G.....C..C.TT....G...TT..TTA.T.CT.....T....G...C.C.T..T.T...A
Aaran-Kav3 ????????????????????????????????????????????????????????????????????????????????????
Aaran-Mad1 ????????????????????????????????????????????????????????????????????????????????????
Aaran-Mad2 ????????????????????????????????????????????????????????????????????????????????????
Aaran-Mad3 ..C.A...C...TCT.T..G.....C....TT...TG...TT..TTA.T.CT.....T....G...C.C.T..T.T...A
Aaran-Sar1 ..C.A...C...TTT.T..G.....C....TT...TG...TT..TTA.T.CT.....T....G...C.C.T..T.T...A
Aaran-Sar2 ..C.A...C...TCT.T..G.....C....TT...TG...TT..TTA.T.CT.....T....G...C.C.T..T.T...A
Aaran-Sar3 ..C.A...C...TTT.T..G.....C....TT...TG...TT..TTA.T.CT.....T....G...C.C.T..T.T...A
Aarma-Mex1 ????????????????????????????????????????????????????????????????????????????????????
Aarti-Pan1 ..G.....C...AC..CCC..T.........AT.TT....TTC.A.AC.....G.A.T.........C.GG.......C..
Aarti-Pan2 ..G.....C...AC..CCC..T.........AT.TT....TTC.A.AC.....G.A.T.........C.GG.......C..
Aarti-Pan3 ..G.....C...AC..CCC..T.........AT.TT....TTC.A.AC.....G.A.T.........C.GG.......C..
Aarti-SCa1 ????????????????????????????????????????????????????????????????????????????????????
Aarti-SCa2 ????????????????????????????????????????????????????????????????????????????????????
Abisp-Cor1 ????????????????????????????????????????????????????????????????????????????????????
Abisp-Cre1 ????????????????????????????????????????????????????????????????????????????????????
Abisp-Cre2 ????????????????????????????????????????????????????????????????????????????????????
Abisp-Cre3 C.TT....C....C.TCC.CTT...CTT...T...TT.C.T.C.ACCCTCGT.T.A.A.TA...GT....T.........
Abisp-Her1 ????????????????????????????????????????????????????????????????????????????????????
Abisp-Her2 ????????????????????????????????????????????????????????????????????????????????????
Abisp-Her3 ????????????????????????????????????????????????????????????????????????????????????
Abisp-Sar1 C.TT....C....C.TCC.CT....CTT...T...TT.C.T.C.ACCCTCGT.T.A.A.TA...GT..............
Abisp-Sar2 C.TT....C....C.TCC.CT....CTT...T...TT.C.T.C.ACCCTCGT.T.A.A.TA...GT...G..........
Abisp-Sar3 C.TT....C....C.TCC.CTT...CTT...T...TT.C.T.C.ACCCTCGT.T.A.A.TA...GT....T.........
Acing-Pan1 C.......C....T..C.G..T..G.....TA....G...TT..C.CC.CCG....GT.TA.G...C...T....G.G..
Acomp-Flo1 ????????????????????????????????????????????????????????????????????????????????????
Acomp-Flo2 ????????????????????????????????????????????????????????????????????????????????????
Acomp-Flo3 ????????????????????????????????????????????????????????????????????????????????????
Adupl-Bim1 C.......C...AT..TCC..TG.......TA...T....TTC.ATACT......G..T.........G......T.C..
Adupl-Bim2 C.......C...AT..TCC..TG.......TA...T....TTC.ATACT......G..T.........G......T.C..
Adupl-Flo1 ????????????????????????????????????????????????????????????????????????????????????
Adupl-SCa1 ????????????????????????????????????????????????????????????????????????????????????
Adupl-SCa2 ????????????????????????????????????????????????????????????????????????????????????
Aerin-Pan1 ....A...C...........T.........T........TTT..A........G.................T.......
Agran-Aus1 C..TC...C....T..T.....T..C.T..TA.T.T..T..TC.AT.CTTAG.C.A.TT.C....T...G...C...C..
Aindi-Bru1 C.G.........A.....G.T..A.CTT..TA........T...ATAAA.AGT.....TTA.T...C.C...T......A
Aindi-Pak1 C..G........A..G....TT.AGCTT..TATG......T...ATAAA.A.T.....TTA.T...C.C......G...A
Aindi-Pak2 C..G........A..G....TT.AGCTT..TATG......T...ATAAA.A.T.....TTA.T...C.C......G...A
Aindi-Tha1 C.GGC.......A.......TT.C.C...GTA.G......T...ATAAT.A.T..A..TTA.T..TC.C.........A
Airre-Far1 C..CCT..T....C.GT....T..GT....TATT..G.C.TTC.AT.CTCG..C.AGT....G.......T.........
Airre-Far2 C..CCT..T....C.GT....T..GT....TATT..G.C.TTC.AT.CTCGG.C.AGT....G.......T.........
Airre-Far3 C..CCT..T....C.GT....T..GT....TATT..G.C.TTC.AT.CTCGG.C.AGT....G.......T.........
Airre-Kav1 C.GTC...T....T..T.....T..C.T..TA.T.TG.T.TTC.AT.CTCA..C.A..T......T..G....CT..A..
Airre-Kav2 C.GTC...T....T..T.....T..C.T..TA.T.TG.T.TTC.AT.CTCA..C.A..T......T..G....CT..A..
Airre-Kav3 ..TC.T..T....CTGCC..T.T....T.CTATC....C.T.A.AT.....T.CCA.........TCC.....T.T.A..
Airre-Mad1 C.GTC...T....T..T.....T..C.T..TA.T.TG.T.TTC.AT.CTCA..C.A..T......T..G....CT..A..
Airre-Nor1 ????????????????????????????????????????????????????????????????????????????????????
Airre-Sar1 ??????????????..T.....T..C....TA.T.T..T..TC.AT.CTTAG.C.A.TT.C.GA.T....??????????
Airre-Sar2 C.GTC...T....T..T.....T..C.T..TA.T.TG.T.TTC.AT.CTCA..C.A..T......T..G....CT..A..
Airre-Sar3 C..TC...C....T..T.....T..C.T..TA.T.T..T..TC.AT.CTTAG.C.A.TT.C....T...G...C....C..
Ajava-Bru1 ..TTA...C...TC..C..TT.........T.TT......TTC.A..CA.C.TTCT....C.....C.........T.G.T
Ajons-Sar1 ..TC.T..T....CTGC...T.T....T.CTATC....C.T.A.AT.....T.CCA.........TCC.....T.T.A..
Ajons-Sar2 ..TC.T..T....CTGC...T.T....T.CTATC....C.T.A.AT.....T.CCA.........TCC.....T.T.A..
Ajons-Sar3 ..TC.T..T....CTGC...T.T....T.CTATC....C.T.A.AT.....T.CCA.........TCC.....T.T.A..
Ajons-Sar4 ..TC.T..T....TTGCC..T.T....T.CTATC....C.T.A.AT.....T.CCA.........TCC.....T.T.A..
Ajons-Tso1 ????????????????????????????????????????????????????????????????????????????????????
Ajons-Tso2 ..TC.T..T....CTGCC..T.T....T.CTATC....C.T.A.AT.....T.CCA.........TCC.....T.T.A..
Ajons-Tso3 ..TC.T..T....CTGCC..T.T....T.CTATC....C.T.A.AT.....T.CCA.........TCC.....T.T.A..
Alate-Jap1 C.T.C...C...AT.T....TT....T.C.TT..TT....TAC.A.ACT.CG.T.T...TG....T........T..A..
Amarg-Col1 C.T.C...C....C..TC.........T.G.A....G..TTTC.A.CCA.C..............AC..G..........
```

```
Amarg-Pan1 C.T.C...C....C..TC........TT...A....G..TTTC.A.CCA.C..G...........AC..G.........
Amarg-Pan2 ????????????????????????????????????????????????????????????????????????????????
Amona-Aus1 C.T.C.......TC.C.C.............ATTTT..C.TTC.A.A.G....G.T.TTTA....T......TC.G....
Aniti-Col1 C.GG....T...TC.G.C..T.......CG.A.T.T.....TC.A..C...T........C......GG....CT..A..
Aplat-Cyp1 ????????????????????????????????????????????????????????????????????????????????
Aplat-Cyp2 C.TC....C....C..C..GTT...CT.C.TT....T.C.TTC.A..C.TGT.C.A.A.TA.........T.........
Aplat-Cyp3 C.TC....C....C..C..GTT...CT.C.TT....T.C.TTC.A..C.TGT.C.A.A.TA.........T.T.......
Aplat-LaH1 ????????????????????????????????????????????????????????????????????????????????
Aplat-Sar1 C.TC....C....C..C..GTT...CT.C.TT....T.C.TTC.A..C.TGT.C.A.A.TA...................
Aplat-Sar2 C.TC....C....C..C..GTT...CT.C.TT....T.C.TTC.A..C.TGT.C.A.A.TA...................
Aplat-Sar3 C.TC....C....C..C..GTT...CT.C.TT....T.C.TTC.A..C.TGT.C.A.A.TA...................
Aplat-Tou1 ????????????????????????????????????????????????????????????????????????????????
Apoly-Fij1 ????????????????????????????????????????????????????????????????????????????????
Apoly-Jap1 C.C.....T...TT...C...T.A...T.TAAT..A...TTTATA.A.AGAGC..TGT...........GT.T..T..G.
Apoly-Jap2 C.C.....T...TT...C...T.A...T.TAAT..A...TTTATA.A.AGAGC..TGT...........GT.T..T..G.
Apoly-Mau1 ????????????????????????????????????????????????????????????????????????????????
Apoly-Mau2 C.TT....T...TT..TC...TGAGG.T.TGAT..A......GTA.A.A.AGT.CTGT..............T.T....
Apoly-NZ1  C.T.G...T...TC.GC......A...T.CTAT.T......TA.A.AAA.G.CC.T.T........C..G.....T.C..
Apoly-NZ2  C.T.G...T...TC.GC......A...T.CTAT.T......TA.A.AAA.G.CC.T.T........C..G.....T.C..
Apoly-Phi1 C..GCT..T....C......TT.A.T.T.TTAT..AG.....A.A.ACA.AGCC...T....T...C.G.T....T..G.
Apoly-Sey1 ...TCT..T...TCTG....TT.G.TTT.TTAT..AG.....A.A.ACA.AGTC...T...GT...C.G.T....T..G.
Apoly-Sey2 ...TCT..T...TCTG....TT.G.TTT.TTAT..AG.....A.A.ACA.A.TC...T...GT...C.G.T....T..G.
Arega-Pan1 .T..N...C...........T...G......A...TG....T..A.C...C......T..C...........T.......
Ascop-Jap1 ..G.CT..C....T..C....T.G.....CT...T.....TTACA.A.A.A.TG.T.T....G.......T....C....
Aside-Pan1 ............................C.TA.G.......T......T.C....T.......................
Asp1-Fij1  ????????????????????????????????????????????????????????????????????????????????
Asp2-Fij1  ..C.CT.......CT.T..T...CG.T...TA..T...TTT..GACA.A.A.TT.T......T...C..GT.........
Asp3-Fij1  CT...T..C....C..T......C......TAT.TT.....T..AT..A.C.TC.T.....GT..T......A....G..
Asp3-Fij2  CT.G.T..C....C..T......C......TAT.TT.N...T..AT..A.C.TC.T.....GT..T......A....G..
Asp3-Fij3  CT...T..C....N..T......C......TAT.NN.A...T..AT..A.C.TC.T.....GT..T......A....G..
Asp4-Fij1  CT..A.AA...AAT.........AGG....TA..TAG..TT.C.CCCC..A.C.......C.....CG.G.....GTA..
Asp4-Fij2  ????????????????????????????????????????????????????????????????????????????????
Asp5-Ton1  .T..C.....TAAC..C......A......TA..TN.N.TT.C.AC....C.T.......A.....C.....A...TG..
Asp6-Phi1  C.T.....N....C.TT...T......T..T..T....C..TC.A.AA..C.C..T.T.TA.....C.....T.......
Asp7-Aus1  C.TCA...T...T...C..CTT.........T.TT......T.C.A..CA..GTT.T..A.GG...TC.........T.G..
Aspin-Cor1 CTGTC...C....T..TC....TC.C....TA.C.TG.TTTT..AC..TCG..CCG.T....G..........C.G....
Aspin-Cre1 CTGTC...C....T..TC....TC.C....TA.C.TG.TTTT..AC..TCG..CCG.T....G..........C.G....
Atris-Oah1 ????????????????????????????????????????????????????????????????????????????????
Avapp-Bru1 C.T.C...T....C.........C...T.CTAT.T.G....TA.A.AC..A.TC.T.TT.C.....C.........T.C.T
Averr-Cal1 ......................G........A.........T.....................................
Averr-Cal2 .T......C......G............C.T..G..G...TT..AT..T.CG...T................TT...CG.
Averr-Cal3 .T......C......G.........G..C.T..G..G...TT..AT..T..G...T................TT...CG.
Averr-Pan1 C.......C.......T...TT.........T.....G....T..A..............C...........T.....G.
Azebr-PNG1 C..TC...C....T..T.....T..C.T..TA.T.T..T..TC.AT.CTTAN.C.A.TT.C....T...G...C...C..
Astro-Naz1 CT......C..AA.TC.C....GA.C....AAT.....C.TG.TATACAG.T...A.TA.A....AACG.TG..T....A
Astro-Ton1 CT......C..AA.TT.C.......C.T..AA......C.TA.TATACA..T.G.A.TA.A....AA...TGT.T....A
Cphor-Mar1 ????????????????????????????????????????????????????????????????????????????????
Cpleu-Fij1 ????????????????????????????????????????????????????????????????????????????????
Pseud-Pan1 C.TTA...C...A.TTCC.......T.TC.AA.T.TG.T..A.CAT.CA.A.T..A.AA.A....AAT..T.TC.C.C..
Tpeni-SDi1 C.C.C....CTAA.T.TC....T..C...TAA.TTT...TT.A.A..AA...C..A.A......TTCC..T..C.TTA..
Tsubi-Far1 .TGGA...C...A...CC..TT.C.C.T..AAT.TTC.T.TA.TATC.T.A....A..G......AA...T.TT.....T
```

```
[          1111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111]
[          1111111111111111111111111111111111112222222222222222222222222222222222222233333333333]
[          2222333344455555666667778888999999000011112233334444555666777788899999000000011111]
[          1369256814703469256891470369023568147036958147803492581470369269023581457890136 9]
Aoers-Pan1 CAACGCCCCATAACTTGGATCAAAAATCGTGTCTCTCCCTACTCCTCTAAGAACAATAGCTTGCCGCTCAACCTCTTTCC
Aafri-STo1 T.TTC.TAT.AG..A.AA.AT..C.....G.....C...C.T........A..T.T..CTC.AT.A...CC.........
Aalli-Col1 ????????????????????????????????????????????????????????????????????????????????????????????
Aamer-Flo1 ????????????????????????????????????????????????????????????????????????????????????????????
Aanti-Col1 .TG.A..ATC..G.ACAA..T.....C..A..............C.......T....A.C.AT.A....GTG....C.T
Aanti-Pan1 .TGTA.TATC..GTACAA..T.....C..A..............C.......T....A.C.AT.A....GTT..G.C.T
Aaran-Cre1 ...T.TTGT...GTACAAG.TG..T..T...ATA.C..T...C...TA.....TC...CT.CAT.C.ATTG.T..A..T.
Aaran-Cre2 ...T.TTGT...GTACAAG.TG..T..T...ATA.C..T...C....A.....TCC..CT.CAT.C.ATTG.T..A..T.
Aaran-Cre3 ...T.TTGT...GTACAAG.TG..T..T...ATA.C..T...C....A.....TCC..CT.CAT.C.ATTG.T..A..T.
Aaran-Far1 ...T.TTGT...GTACAAG.TG..T..T...ATA.C..T...C....A.....TCC.GCT.CAT.C.ATTG.T..A..T.
Aaran-Far2 ...T.TTGT...GTACAAG.TG..T..T...ATA.C..T...C....A.....TCC..CT.CAT.C.ATTG.T..A..T.
Aaran-Far3 ...T.TTGT...GTACAAG.TG..T..T...ATA.C..T...C....A.....TCC..CT.CAT.C.ATTG.T..A..T.
Aaran-Kav1 ...T.TTGT...GTACAAG.TG..T..T...ATA.C..T...C....A.....TCC..CT.CAT.C.ATTG.T..A..T.
Aaran-Kav2 ...T.TTGT...GTACAAG.TG..T..T...ATA.C..T...C...TA.....TCC..CT.CAT.C.ATTG.T..A..T.
Aaran-Kav3 ????????????????????????????????????????????????????????????????????????????????????????????
Aaran-Mad1 ????????????????????????????????????????????????????????????????????????????????????????????
Aaran-Mad2 ????????????????????????????????????????????????????????????????????????????????????????????
Aaran-Mad3 ...T.TTGT...GTACAAG.TG..T..T...ATA.C..T...C....A.....TCC..CT.CAT.C.ATTG.T..A..T.
Aaran-Sar1 ...T.TTGT...GTACAAG.TG..T..T...ATA.C..T...C....A.....TCC..CT.CAT.C.ATTG.T..A..T.
Aaran-Sar2 ...T.TTGT...GTACAAG.TG..T..T...ATA.C..T...C....A.....TCC..CT.CAT.C.ATTG.T..A..T.
Aaran-Sar3 ...T.TTGT...GTACAAG.TG..T..T...ATA.C..T...C....A.....TCC..CT.CAT.C.ATTG.T..A..T.
Aarma-Mex1 ????????????????????????????????????????????????????????????????????????????????????????????
Aarti-Pan1 .TGTA..ATC..GTACAA..T.....C..A...........T...C...........A.C.AT.A....G.T..A.C.T
Aarti-Pan2 .TGTA..ATC..GTACAA..T.....C..A...........T...C...........A.C.ATTA....G.T..A.C.T
Aarti-Pan3 .TGTA..ATC..GTACAA..T.....C..A...........T...C...........A.C.AT.A....G.T..A.C.T
Aarti-SCa1 ????????????????????????????????????????????????????????????????????????????????????????????
Aarti-SCa2 ????????????????????????????????????????????????????????????????????????????????????????????
Abisp-Cor1 ????????????????????????????????????????????????????????????????????????????????????????????
Abisp-Cre1 ????????????????????????????????????????????????????????????????????????????????????????????
Abisp-Cre2 ????????????????????????????????????????????????????????????????????????????????????????????
Abisp-Cre3 ...T..TAT...GTACAA.C.T..G.CT.G.A.A.CA...GT...C..??????????????????????????????????
Abisp-Her1 ????????????????????????????????????????????????????????????????????????????????????????????
Abisp-Her2 ????????????????????????????????????????????????????????????????????????????????????????????
Abisp-Her3 ????????????????????????????????????????????????????????????????????????????????????????????
Abisp-Sar1 ...T..TAT....TAC.A.C.T..C.CT.A.A.A.CA...GT...C.........T..T.CCAT.T...TC.T....C.T
Abisp-Sar2 ...T..TAT....TAC.A.C.T..C.CT.G.A.A.CA...GT...C.........T..T.CCAT.T.C.TC.T....C.T
Abisp-Sar3 ...T..TAT...GTACAA.C.T..G.CT.G.A.A.CA...GT...C.........T..T.CC.T.C.C.T??????????
Acing-Pan1 T.TTAT..T.A...ACAAGAT..C...........C......C.T.....C.CT....A.CCAT.TTAT...T..C.C.A
Acomp-Flo1 ????????????????????????????????????????????????????????????????????????????????????????????
Acomp-Flo2 ????????????????????????????????????????????????????????????????????????????????????????????
Acomp-Flo3 ????????????????????????????????????????????????????????????????????????????????????????????
Adupl-Bim1 .TGTAT..TC...TACAA..T.....C..A..............C....A..T.....TCCAT.C......T..A.C.T
Adupl-Bim2 .CGTAT..TC...TACAA..T.....C..A..............C....A..T....ATCCAT.C......T..A.C.T
Adupl-Flo1 ????????????????????????????????????????????????????????????????????????????????????????????
Adupl-SCa1 ????????????????????????????????????????????????????????????????????????????????????????????
Adupl-SCa2 ????????????????????????????????????????????????????????????????????????????????????????????
Aerin-Pan1 ............C.CAAA........AT.......C.T...TC.......A..T....AT..A....C.G..T..C.C..
Agran-Aus1 T...ATTAT.....ACA.....TC.....A......TT...T..TC.......T.TC.C.C.A..A.CT.GTG.......
Aindi-Bru1 T.GTA.TATCA...AAAA.C.TCC.CA....A.C.CT.TC.......AT....T....C.A.AT.C.ATGG.T..A...A
Aindi-Pak1 T.GT..TA.CA...AAAA...CCT.CA....A.C.CTT...T.....AC....T....C.ACAT.C.AT.G.T..A...A
Aindi-Pak2 T.GT..TATCA...AAAA...CCT.CA....A.C.CTT...T.....AC....T....C.ACAT.C.AT.G.T..A...A
Aindi-Tha1 T..TA.TATCA...AA.A...CCCC.A........C......C..C.AT....T....C.G.A..C.AT......A.C.A
Airre-Far1 T..TA.TAT....TACAA..T..T..C.......T.T.....CT.........T.T..TTCC.T.T.C....G......A
Airre-Far2 T..TA.TAT....TACAA..T..T..C.......T.T.....CT.........T.T..TTCC.T.T.C....G......A
Airre-Far3 T..TA.TAT....TACAA..T..T..C.......T.T.....CT.........T.T..TTCC.T.T.C....G......A
Airre-Kav1 T...A.TAT.....ACAA.C.GTC.....G.....CTT..G...TC.......TGTC.C.CCA..T.CT.GTG....C.T
Airre-Kav2 T...A.TAT.....ACAA.C.GTC.....G.....CTT..G...TC.......TGTC.C.CC...T.CT.GTG....C..
Airre-Kav3 T..TA.TAT..T..ACTAGAT.TC..CT.....A.C..T.....TG.CG....T.T.TC.CCA..C...TT.........
Airre-Mad1 T...A.TAT.....ACAA.C.GTC.....G.....CTT..G...TC.......TGTC.C.CCA..T.CT.GTG....C.T
Airre-Nor1 ????????????????????????????????????????????????????????????????????????????????????????????
Airre-Sar1 ????????????????????????????????????????????????????????????????????????????????????????????
Airre-Sar2 T.G.A.TAT.....ACAA.C.GTC.....G.....CTT..G...TC....A..TGTC.C.CC...T.CT.GTG..?.C..
Airre-Sar3 T...ATTAT..G..ACA.....TC.....A......TT...T..TC.......T.TC.C.C.A..A.CT.GTG.......
Ajava-Bru1 T.T.ATTAT..T..A.AC...CCCTTCT...A....T.......T..A..AGCT.T.TCT.CAT.T.CT...T....C.A
Ajons-Sar1 T..TA.TAT..T..ACTAGAT.TC..CT.....A.C..T.....TG.CG....T.T.TC.CCA..C...TT.........
Ajons-Sar2 T..TA.TAT..T..ACTAGAT.TC..CT.....A.C..T.....TG.CG....T.T.TC.CCA..C...TT.T.......
Ajons-Sar3 T..TA.TAT..T..ACTAGAT.TC..CT.....A.C..T.....TG.CG....T.T.TC.CCA..C...TT.........
Ajons-Sar4 T..TA.TAT..T..ACTAGAT.TC..CT.....A.C..T.....TG.CG....T.T.TC.CCA..C...TT.........
Ajons-Tso1 ????????????????????????????????????????????????????????????????????????????????????????????
Ajons-Tso2 T..TA.TAT..T..ACTAGAT.TC..CT.....A.C..T.....TG.CG....T.T.TC.CCA..C...TT.........
Ajons-Tso3 T..TA.TAT..T..ACTAGAT.TC..CT.....A.C..T.....TG.CG....T.T.TC.CCA..C...TT.........
Alate-Jap1 .GGTAT.A..A...ACAA.C..CC.C.T...A.C.CT.T..TCTTCTA..A.GT.T..CT..A..C.GT........C..
Amarg-Col1 ....AT......N.ACAA..T..........C..T.T.....CTTC.......T.C..A.ACAT.A.AT...T..A..TA
```

```
Amarg-Pan1 ....A.......GTACAA..TG.........C..T.T.....CTTC.C.....T.C..A.ACAT.A.AT...T..A..TA
Amarg-Pan2 ?????????...G.ACAA..T..........C..T.T.....CTTC.......T.C..A.ACAT.A.AT...T..A..TA
Amona-Aus1 ..CTA...T.A..TACAA..TGTT.CCT...C.C.CT..A..C...TA..C....TC.C.CCA....A.......C..T.
Aniti-Col1 ...TAT..T..C..ACAA.C.G.......C.C..........C.T........T.C..ATCC.T.A.C....T..A....
Aplat-Cyp1 ??????????????????????????????????????????????????????????????????????????????
Aplat-Cyp2 T.G.AT.ATC....AC.A.CTG.....T...A.C.C..TAG....C....A..T.T..T.CC.T.C.C.CT........T
Aplat-Cyp3 T.G.AT.ATC....AC.A.CTG.....T...A.C.C..TAG....C....A..T.T..T..C.T.C.C.CT........T
Aplat-LaH1 ??????????????????????????????????????????????????????????????????????????????
Aplat-Sar1 T.G.AT.ATC..G.AC.A.CTG.....T...A.C.C..TA.....C....A..T.T..T.CC.T.C.C.T??????????
Aplat-Sar2 T.G.AT.ATC..G.AC.A.CTG.....T...A.C.C..TA.....C....A..T.T..T.CC.T.C.C.T??????????
Aplat-Sar3 T.G.AT.ATC..G.AC.A.CTG.....T...A.C.C..TA.....C....A..T.T..T.CC.T.C.C.T??????????
Aplat-Tou1 ??????????????????????????????????????????????????????????????????????????????
Apoly-Fij1 ??????????????????????????????????????????????????????????????????????????????
Apoly-Jap1 T.TT.TTA..A..TA.TA.AT.CC...T.....A.C..T.....T.....A..T.CCGC.CCAT.C.CTC..T....C.A
Apoly-Jap2 T.TT.TTA..A..TA.TA.AT.CC...T.....A.C..T.....T.....A..T.CCGC.CCAT.C.CTC..T....C.A
Apoly-Mau1 ??????????????????????????????????????????????????????????????????????????????
Apoly-Mau2 T.TT.TTA..A..TA.TA.ATGCC...T.....ATC.TT.............T.C..C..CAT.C.CTC..T....C.A
Apoly-NZ1  TG..AT.G...G.TA.AAGA..CCT....C.C.C.C..T..TC.TA....TG...T..CT.CAT.T.CTTG....C.C.A
Apoly-NZ2  TG..AT.G...G.TA.AAGA..CCT....C.C.C.C..T..T..TA....TG...T..CT.CAT.T.CTTG....C.C.A
Apoly-Phi1 ..TTA..AT.A..TA.TA.AT.CC...T...C......T.....T..C.....T.T..C..CA..T.ATC..T......A
Apoly-Sey1 T.TTA..AT.A..TACTA.AT.CC..CT...C.C....T.....T..C.....T.T..C.CC...T.ATCG.T......A
Apoly-Sey2 T.TTA..AT.A..TA.TA.AT.CC..CT...C.C....T.....T..C.....T.T..C.CC...T.ATCG.T......A
Arega-Pan1 .GGT....T.CGC.C..A..T.....A........C.T...T......G....A....AT.C.T.A....G.T..C...T
Ascop-Jap1 T..TA..AT.A..TA.TA.C..CC...T...C.C..T.T...CTTC....A..T.C.GTTCCAT.CTGTCG.T....C.A
Aside-Pan1 ...TA...T...G....A...G.............C..............T....A.............T..C....
Asp1-Fij1  ??????????????????????????????????????????????????????????????????????????????
Asp2-Fij1  T.C.A..AT..GC.ACAA.CT..C.......C.A.CT.....C.TA.........C..CTCCAT.C.AT.G.A..A.C..
Asp3-Fij1  T.CTATTAT.AG..ACAA.A...C.C.....A.AT.....GT..T..CT.A.......TT.CA..T.CT...T..A.C.A
Asp3-Fij2  T.CTAT.AT.AG..ACAA.A...C.C.....A.AT.....GT..T..CT.........TT.CA..T.CT...T..A.C.A
Asp3-Fij3  T.CTATTAT.AG..ACAA.A...C.C.....A.AT.....GT..T..CT.A.......TT.CA..T.CT...T..A.C.A
Asp4-Fij1  T.CTA.TAT.AC..ACCA.....C.T..A.AA.GT.....T.CTT..CC.A.....C.TT.CAT.C.GTGG....A...A
Asp4-Fij2  ??????????????????????????????????????????????????????????????????????????????
Asp5-Ton1  T...A.TAT.AT..ACTA.....T.C..A.AA.A.CTT..T.CTTC.CC.A.....C.TT.CAT.C.AT......A.C.A
Asp6-Phi1  T.G.A.TGT.A..TA..A.AT.CC.CC...A..CTCTT....CT...A.....T.CC.ATC.A..A.AT..TA..A...A
Asp7-Aus1  ..C.ATTAT.....A.AC...CCCCTCT.C.A..T.T.......TCTA.T..C..C.CC...AT.T.AT.C.T....C.G
Aspin-Cor1 T.G.A.TAT......C.A.C.GTC.....G.C.C.C.T..GTC.TA.C..T.TT.T..CTCCAT.T.C..C.G.......
Aspin-Cre1 T.G.A.TAT......C.A.C.GTC.....G.C.C.C.T..GTC.TA.C..T.TT.T..CTCCAT.T.C..C.G.......
Atris-Oah1 ??????????????????????????????????????????????????????????????????????????????
Avapp-Bru1 .G.T.T.AT.C..TA.A..A..C.T........C.C......C..A.........T..CTCCAT.T..TC..T..C.C.A
Averr-Cal1 ....A...T.....................................................T.........T.A.............
Averr-Cal2 ...TAT..........AA..TG....A...A....C.T.....................AT.C...A...GG.T.......
Averr-Cal3 ...TAT..........AA..TG....A...A....C.T.....................AT.C...A...GG.T.......
Averr-Pan1 T...A..A....C.C..A..T.....C........C..T...C.T.....A..T....AT.CA..A...G..T.......
Azebr-PNG1 T...ATTAT.....ACA.....TC.....A......TT...T..TC.......T.TC.C.C.A..A.CT.GTG.......
Astro-Naz1 ....AT.ATCA...GCAT...CCCTT...AA...T....C.TC..C..C..T.T....C.ACAT.A.ATGGGTGG..A.A
Astro-Ton1 ....AT.ATCA...CCAT....CCTT...GA...T...T..T...C..C..T.T....C.ACA..A.AT..GTGG..C.A
Cphor-Mar1 ??????????????????????????????????????????????????????????????????????????????
Cpleu-Fij1 ??????????????????????????????????????????????????????????????????????????????
Pseud-Pan1 T.CTAT.TTCAGT.AAAA.A..TT.TAT.AA...TCT.T...C...TAC.AT.T....A.CCAT.C.AT.GGAGGA..TT
Tpeni-SDi1 ..C.CT.AT....TA.AA...TCC.TAT.AA...T.T...CT.T.CTA....T..T..AT.CA...TATCC.AG..C.TT
Tsubi-Far1 T....T.TTCA.C.AAAT.C..CTTCA..AA..CT.TT..G.C.....C..C.T.C....CCA..C.AT.GGGGGAA...
```

```
[          1111111111111111111111111111111111]
[          3333333333333333333333333333333333]
[          2222223334444444555566666777788888]
[          0245682470123456235814570369257 89]
Aoers-Pan1 CCGAGAAACTATTCTTTAAGTCCATTCAACTCC
Aafri-STo1 .A..A..TTCG....A...TA..GAC..G..TT
Aalli-Col1 ??????????????????????????????????
Aamer-Flo1 ??????????????????????????????????
Aanti-Col1 ....A..C.C..C..A.GT..T???????????
Aanti-Pan1 ...GA..C.C..C..A.GT..T..ACT....T.
Aaran-Cre1 ....A..T..GCCT.AC...G.TGC.T.G....
Aaran-Cre2 ....A..T..GCCT.AC...G.TGC.T.G....
Aaran-Cre3 ....A..T..GCCT.AC...G.TGC.T.G....
Aaran-Far1 ....A..T..GCCT.AC...G.TGC.T.G....
Aaran-Far2 ....A..T..GCCT.AC...G.TGC.T.G....
Aaran-Far3 ....A..T..GCCT.AC...G.TGC.T.G....
Aaran-Kav1 ....A..T..GCCT.AC...G.TGC.T.G....
Aaran-Kav2 ....A..T..GCCT.AC...G.TGC.T.G....
Aaran-Kav3 ??????????????????????????????????
Aaran-Mad1 ??????????????????????????????????
Aaran-Mad2 ??????????????????????????????????
Aaran-Mad3 ....A..T..GCCT.AC...G.TGC.T.G....
Aaran-Sar1 ....A..T..GCCT.AC...G.TGC.T.G....
Aaran-Sar2 ....A..T..GCCT.AC...G.TGC.T.G....
Aaran-Sar3 ....A..T..GCCT.AC...G.TGC.T.G....
Aarma-Mex1 ??????????????????????????????????
Aarti-Pan1 ....A..C.C..C..A..T..T..ACT....T.
Aarti-Pan2 ....A..C.C..C..A..T..T..ACT....T.
Aarti-Pan3 ....A..C.C..C..A..T..T..ACT....T.
Aarti-SCa1 ??????????????????????????????????
Aarti-SCa2 ??????????????????????????????????
Abisp-Cor1 ??????????????????????????????????
Abisp-Cre1 ??????????????????????????????????
Abisp-Cre2 ??????????????????????????????????
Abisp-Cre3 ??????????????????????????????????
Abisp-Her1 ??????????????????????????????????
Abisp-Her2 ??????????????????????????????????
Abisp-Her3 ??????????????????????????????????
Abisp-Sar1 ....A..C..G....A...CAAT.A.T.G...T
Abisp-Sar2 ....A..C..G....A...CAAT.A.T.G...T
Abisp-Sar3 ??????????????????????????????????
Acing-Pan1 .....GGT.C..C..ACG......C.???????
Acomp-Flo1 ??????????????????????????????????
Acomp-Flo2 ??????????????????????????????????
Acomp-Flo3 ??????????????????????????????????
Adupl-Bim1 .T.....T.C..C..ACG......ACT....T.
Adupl-Bim2 .T.....T.C..C..ACG......ACT....T.
Adupl-Flo1 ??????????????????????????????????
Adupl-SCa1 ??????????????????????????????????
Adupl-SCa2 ??????????????????????????????????
Aerin-Pan1 .....GGG.C..C.........T..C...T...
Agran-Aus1 ....A..TT...A......CA.T.A...G..T.
Aindi-Bru1 ...GA..TT.GC...A..TA.AT..C.......
Aindi-Pak1 .T..A..TTCG.C..A..CA.AT..C...T.T.
Aindi-Pak2 .T..A..TTCG.C..A..CA.AT..C...T.T.
Aindi-Tha1 .T..A..T..G..T.A..T..AT.AC...T...
Airre-Far1 ....AGGT.C..C......TA...A.T.G..T.
Airre-Far2 ....A.GT.C..C......TA...A.T.G..T.
Airre-Far3 ....A.GT.C..C......TA...A.T.G..T.
Airre-Kav1 ....A.GT....A......CA.T.A.T....T.
Airre-Kav2 ....A.GT....A......CA.T.A.T.G..T.
Airre-Kav3 ....A.GC.C.....A..TAC.T..CTT.T...
Airre-Mad1 ....A.GT....A......CA.T.A.T....T.
Airre-Nor1 ??????????????????????????????????
Airre-Sar1 ??????????????????????????????????
Airre-Sar2 ....A.GT.C..A.A....CA...A.???????
Airre-Sar3 ....A.GTT...A......CA.T.A.T.G..T.
Ajava-Bru1 ....A..T..GCC..A..TAAAT..CTC.....
Ajons-Sar1 ....A.GC.C.....A..TAC.T..CTC.T...
Ajons-Sar2 ....A.GC.C.....A..TAC.T..CTC.T...
Ajons-Sar3 ....A.GC.C.....A..TAC.T..CTC.T...
Ajons-Sar4 ....A.GC.C.....A..TAC.T..CTT.T...
Ajons-Tso1 ??????????????????????????????????
Ajons-Tso2 ....A.GC.C.....A..TAC.T..CTT.T...
Ajons-Tso3 ....A.GC.C.....A..TAC.T..CTT.T...
Alate-Jap1 ....A.G..C......C.T..A..C......T.
Amarg-Col1 .T.N.G.C.......A.....T..C.TC.....
```

```
Amarg-Pan1 .T...G.C.......A.....T..?????????
Amarg-Pan2 .T.G.G.C.......A.....T..C.TC.....
Amona-Aus1 .T..A..C.....T.A..TTAA...C.....T.
Aniti-Col1 .......C.C..C..AC.T.?????????????
Aplat-Cyp1 ??????????????????????????????????
Aplat-Cyp2 .T..AG.T..G....A...TA...?????????
Aplat-Cyp3 .T..AG.T..G....A???????????????????
Aplat-LaH1 ??????????????????????????????????
Aplat-Sar1 ??????????????????????????????????
Aplat-Sar2 ??????????????????????????????????
Aplat-Sar3 ??????????????????????????????????
Aplat-Tou1 ??????????????????????????????????
Apoly-Fij1 ??????????????????????????????????
Apoly-Jap1 .T..A..CT...C..A..TAAAT..C.G.....
Apoly-Jap2 .T..A..CT...C..A..TAAAT..C.G.....
Apoly-Mau1 ??????????????????????????????????
Apoly-Mau2 .T..A..C..GCC..G.GT..AT..C.GG....
Apoly-NZ1  ....A..C..GCCT.A....AA...C.......
Apoly-NZ2  ....A..C..GCCT.A....AA...C.......
Apoly-Phi1 ....A.....GCA..A.GT..A....T.G....
Apoly-Sey1 ....A.....GCA..A..T..A....T......
Apoly-Sey2 ....A.....GCA..A..T..A....T......
Arega-Pan1 TA.G..GT.????????????????????????
Ascop-Jap1 .T..A..C.CGCC..A.G.TGA.C.C.C.....
Aside-Pan1 ...G.....C..C........T.......T...
Asp1-Fij1  ??????????????????????????????????
Asp2-Fij1  .A...G.T..GCAT.A..TA.A...CT....T.
Asp3-Fij1  .A..A....C..A..A..TAAA..CCTC.....
Asp3-Fij2  ....A....C..A..A..TAAA..CCTC.....
Asp3-Fij3  .A..A....C..A..A..TAAA..CCTC.....
Asp4-Fij1  ....A..C....A..A..T.AA..C.TT.T...
Asp4-Fij2  ??????????????????????????????????
Asp5-Ton1  ....A..C.C..A..AC.CAAA..C.TC...T.
Asp6-Phi1  .T.GA..T.......A..TTAA...CT......
Asp7-Aus1  .T..A..T..GGC..A..TAA???????????
Aspin-Cor1 ....A.GC....A....G.CA.T.A.T....T.
Aspin-Cre1 ....A.GC....A....G.CA.T.A.T....T.
Atris-Oah1 ??????????????????????????????????
Avapp-Bru1 .????????????????????????????????
Averr-Cal1 ......G..................T..TA??
Averr-Cal2 .T....G..C..C...ACTTG.....T...
Averr-Cal3 .T....G..C..C...ACTTG.....T...
Averr-Pan1 .....G...C..C...C.T......CT......
Azebr-PNG1 ....A.GTT...A......CA.T.A.T.G..T.
Astro-Naz1 .A..A.GC....C.....T.....A...G..T.
Astro-Ton1 .A..A.GC....C.....T.....A.T....T.
Cphor-Mar1 ??????????????????????????????????
Cpleu-Fij1 ??????????????????????????????????
Pseud-Pan1 TA.....TT......C..T........T.T.TT
Tpeni-SDi1 ....A..C.C...T.A..T.A..T..TC....T
Tsubi-Far1 TAT....CTC.....A..T.....C..T...T.
```

CHAPTER II

Reconstruction of the evolutionary history in sea stars of the genus *Astropecten* (Asteroidea:Paxillosida:Astropectinidae) using molecular markers, fossil records and geological events

Deborah E. Zulliger

Abstract

In theory, taxa with long planktotrophic larval stages should rarely speciate due to high effective population sizes and abundant gene flow. Although this assumption has been questioned by some studies, the species diversity in high dispersal sea stars of the genus *Astropecten* is nevertheless remarkable. This raises questions about the evolutionary history of this group, the study of which is generally seen as a window to past speciation patterns and processes. The present study therefore aimed at reconstructing the evolutionary history of *Astropecten* lineages by dating a molecular phylogeny using geological events and fossil records. Using the same mtDNA sequence data which was produced and presented in Chapter I and after removing redundant taxa, I inferred a molecular phylogeny and calibrated the resulting tree using a) geological events, b) fossil records and c) a combination of geological events and fossil records. Divergence times of extant *Astropecten* lineages were estimated using a relaxed clock method as implemented in the software MULTIDIVTIME, and results from the different calibration categories were then compared. According to my results, recent lineages began to diverge in the Mid-Miocene, whereas the species diversity in the East Pacific and West Atlantic were most likely enhanced by the rise of the Panama Isthmus in the Late Miocene and Early Pliocene. Furthermore, the majority of exclusively Mediterranean species have most likely evolved after the Messinian Salinity Crisis (Late Miocene, ~5 mya), suggesting a rather recent origin in the Mediterranean of these lineages. Ancient lineages, dating back to the Mid-Miocene, can be found in species from the South Pacific and in *A. aranciacus* of the Mediterranean and East Atlantic. Phylogenetic relationships of these lineages, however, could not be satisfactorily resolved with the mtDNA regions used in this study.

Introduction

With over 150 species described in the literature (Say 1825; Gray 1841; Müller and Troschel 1842; Perrier 1875/6; Agassiz 1877; Sladen 1889; Ludwig 1897; Fisher 1906; Koehler 1909; Koehler 1910; Fisher 1911; Fisher 1913; Verrill 1914; Verrill 1915; Döderlein 1917; Koehler 1924; Clark and Downey 1992), the diversity within sea stars of the genus *Astropecten* (Asteroidea:Paxillosida) is remarkable for high dispersal echinoderms. It is not known, however, if all of these species have larvae with extended planktonic stages and to which extent this would lead to a lower speciation rate (see Chapter I). The high species-diversity in *Astropecten* raises questions related to the evolutionary history of this genus as well as the timing of the radiation of this group. One possibility is that *Astropecten* is an ancient group and has therefore slowly generated species by species over a long time period. However, as the results of Chapter I suggest, it is possible that long-lived planktotrophic larvae do not necessarily inhibit speciation as much as has been assumed and that *Astropecten* has recently undergone rapid evolution.

Of particular interest is the species diversity in the Mediterranean, where at least seven *Astropecten* species are currently present. There is no other echinoderm group in this region that is richer in species. Four of these species are endemic to the Mediterranean, whereas the distribution of the other four species extends into the East Atlantic. While the latter may have evolved in either the Mediterranean or in the East Atlantic, it is likely that Mediterranean endemics have also originated in the same basin where they today exclusively are present. An indication for a Mediterranean origin would be if the Mediterranean endemics arose after the Messinian Salinity Crisis (MSC). The Salinity Crisis occurred during the Messinian stage of the Late Miocene, which started ~5.96 mya and resulted in the desiccation of much of the Mediterranean basin (Krijgsman et al. 1999). In the Early Pliocene (~5.33 mya) the MSC ended as the sea level rose and flooded the Mediterranean basin within a few years (Blanc 2002). The MSC had a dramatic impact on marine organisms endemic to the Mediterranean and lead most or even all of them into extinction (Hofrichter 2003).

In contrast, species whose distributions extend beyond the Mediterranean might be older, as they may have persisted in the East Atlantic during the MSC. To determine the origin of these species is difficult, as both an Atlantic and a Mediterranean evolution are plausible.

Investigating the evolutionary history of taxa is essential when studying evolutionary patterns, speciation events and rates of evolution (Smith and Peterson 2002). In order to estimate divergence times of lineages, phylogenetic trees, e.g., inferred from molecular sequence data, can be dated using fossils or geological events. The origin of numerous taxa has been investigated this way, including marine invertebrates such as crustaceans (Perez-Losada et al. 2004; Perez-Losada et al. 2008; Tinn and Oakley 2008), mollusks (Wilke et al. 2000; Williams and Reid 2004; Wood et al. 2007; Frey and Vermeij 2008) and echinoids (Smith et al.

2006). To my knowledge, no studies until now have estimated divergence times in asteroids using molecular data.

An earlier attempt to determine the evolutionary history of the genus *Astropecten* was made by (Döderlein 1917) using morphological criteria. In his monograph, Döderlein stated that the main direction of "species dispersal" occurred from east to west, beginning in the East Atlantic with the most primitive *Astropecten* species and then dispersing into other regions of the globe. The author, however, did not mention fossil records of this group, and no time estimates were suggested for the divergence of *Astropecten* lineages.

In order to obtain a more complete picture of the evolutionary history of *Astropecten*, the present study aimed at applying molecular dating techniques to determine *Astropecten* divergence times by constructing a molecular phylogeny and calibrating it using fossil records and geological events. The molecular phylogeny was derived from mitochondrial DNA (mtDNA) sequences of the 12S rRNA, 16S rRNA and cytochrome oxidase subregion I (COI), which were already produced in Chapter I. I attempted to address several questions related to the evolutionary history of *Astropecten*, such as (1) the approximate divergence time of *Astropecten* lineages, (2) the influence of geological events on lineage divergence, and (3) whether exclusively Mediterranean species diverged before or after the Messinian Salinity Crisis. Furthermore, I explored whether there are differences between divergence time estimates based on independent fossil record and geological event calibrations and how potential inconsistencies could be explained.

Methods

For this study, I used the DNA sequence data and alignment matrix of the three mitochondrial regions 12S rRNA, 16S rRNA and cytochrome oxidase c subregion I (COI) that were produced for 118 *Astropecten* and seven outgroup specimens in Chapter I (see Chapter I for sampling, DNA extraction, PCR and alignment methods). To speed up calculations, redundant taxa were removed according to the results of Chapter I in such a way that only one specimen per species remained in the matrix. The resulting matrix consisted of 53 *Astropecten* and seven outgroup specimens belonging to the genera *Ctenophoraster*, *Ctenopleura*, *Pseudarchaster*, *Tethyaster* and *Thrissacanthias* (Table 1).

Phylogenetic analysis

The program MODELTEST 3.7 (Posada and Crandall 1998) was used to perform hierarchical likelihood ratio tests (hLRTs), which helps to identify the most appropriate nucleotide substitution model for the dataset. MODELTEST selected the general time reversible model with proportion of invariable sites (I) and

gamma distribution (Γ) (GTR+I+Γ) as the best-fit model of evolution (- ln L = 7'989.82; see Figure 1 for parameter information).

In Chapter I we already performed a maximum parsimony and Bayesian analysis, which are adequate methods when using extensive data sets. The reduced data set used in this study, on the other hand, allowed me to perform a more calculation intensive method, the maximum likelihood (ML) analysis. I performed this analysis in PAUP* version 4.01 (Swofford 2003) by heuristic search and 10 random addition replicates applying the TBR branch-swapping option. First, parameters estimated by MODELTEST were used for the ML search. Then, parameters were re-estimated on the resulting tree and used for a subsequent ML search. This process was repeated until the search resulted in the same topology and identical parameter estimates as the previous iteration.

To estimate nodal support, Bayesian analysis was performed in MRBAYES version 3 (Ronquist and Huelsenbeck 2003) assuming the GTR+I+Γ model, unlinking the partitions and setting the chain heating temperature to 0.02. Two runs of four chains were performed for 3'000'000 generations sampling every 500 generations. After plotting the -lnL values to determine the number of generations after which the values become stationary, the first 500 trees (250'000 generations) were discarded as burn-in. From the remaining trees I constructed a consensus tree following the 50% majority rule to estimate posterior probabilities indicating clade credibility, which I then placed on the ML tree.

Likelihood ratio test

To test if the genes evolved with rate constancy across the ML tree as in a molecular clock, I performed a test of rate constancy in PAUP* using a likelihood ratio test (Felsenstein 1988). The molecular clock assumes that mutations accumulate at an approximately constant rate as long as the DNA sequence retains its original functions and extrapolates divergences based on one or more calibration points (Zuckerkandl and Pauling 1965). I obtained likelihood scores in PAUP* for the best fitting model with and without enforcing a molecular clock. The likelihood scores of the two trees were compared, and the difference (times two) was tested for significance in a chi-square table with $n - 2$ degrees of freedom (n = number of terminal taxa in the tree). According to this test, my sequence data rejects the molecular clock hypothesis ($P < 0.005$; enforced molecular clock –ln L = 16'364.19) requiring methods which account for rate variation across branches of the tree to relax the molecular clock when estimating divergence times.

Calibrating the tree

Methods using multiple calibration points often provide more appropriate estimates of divergence times, especially if the assumption of clocklike molecular divergence is violated (Thorne and Kishino 2002; Yang and Yoder 2003). By taking geological events and fossil records into account, I attempted to meet this requirement.

Geological events

The final closure of the Isthmus of Panama is dated at roughly 2.8 – 3.1 mya (Coates and Obando 1996) and led to a complete marine barrier between the tropical Pacific and Atlantic. However, other vicariance events may have lead to an earlier separation of species (Cunningham and Collins 1994; Collins et al. 1996; Knowlton and Weigt 1998; Marko 2002), especially in deep sea organisms (Knowlton et al. 1993), which may have diverged as early as the Mid Miocene. As *Astropecten* are predominantly shallow water sea stars, gene flow is likely to have persisted between the Atlantic and Pacific for longer than in other marine organisms also affected by the rise of the Isthmus. Therefore, I set the separation of the East Pacific clade from the West Atlantic clade (node 98) to a lower limit of 2.8 mya and an upper limit of 6 mya.

In the Mid Miocene ~18 mya, plate collisions began to seal off the Indian Ocean from the Mediterranean, finally closing the Tethyan Sea at the end of the Serravalian ~11.8 mya (Rögl and Steininger 1983; Rögl 1999). This geological event can be used to date the separation of the Indo-West Pacific clade from the East Atlantic and Mediterranean clade (node 104). The date of the node was calibrated by setting the lower limit to 11.8 mya and the upper limit to 18 mya.

Other commonly used geological events used to calibrate phylogenetic trees are the constriction of the Indonesian Seaway about 13 mya (White 1994) and the Benguela cold upwelling, which established about 10 mya (Siesser 1980). These barriers, however, are not applicable to this study, as there is no node that can be attributed to either of them.

Fossil record

Fossils of the genus *Astropecten* have only been found in Cenozoic sediments. Some Oligocene *Astropecten* have been described by Blake (1973) from Oregon in the United States, and more fossil records of *Astropecten* have been found from the Eocene and Miocene (Cavara 1886; Sacco 1893; Del Prato 1896; Kudrin 1957; Heller 1958; Kaczmarska 1987; Borghi 1995; Kroh and Harzhauser 1999; Pereira et al. 2002; Pereira et al. 2003; Rico-García et al. 2008). Although fossils of astropectinid sea stars are abundant, complete articulated records are rare. Following methods devised by Blake (1973), it is possible to assign fossils to the genus level using ossicles, whereas identifications to the species level are doubtful when using these methods. Since many fossil records of *Astropecten* consist only of disintegrated ossicles, it is often not possible to assign them to a species or to determine their relationship to extant species. This is, however, a crucial step when dating internal nodes in a molecular tree. For this reason, I only considered records of articulated fossils which can be confidently assigned to a group or a species of extant *Astropecten*. As an estimate of the first appearance of the genus, I used the oldest fossil record of *Astropecten* sensu stricto, corresponding to *A. granulatus* (Rasmussen 1972) from the late Eocene, Bartonian, UK (ca. 40 million years; replacement name *A. anglicus*; (Nosowska 1997). Older records are doubtful (personal communication A.

Gale), such as *A. postornatus* from the Upper Danian ~62 mya (Rasmussen 1972) and *A. punctatus* and *A. n. sp.* from the Late Maastrichtian and the early Paleocene ~65 mya (Jagt 2000).

In order to calibrate internal nodes of the molecular phylogeny, I considered the following fossils as suitable, as they are articulated and can be clearly aligned with extant species (see also Table 2 for summary): a specimen from the Pliocene (~5.3-1.8 mya) in Italy (Bra), described as *A. bispinosus* by Sacco (1893), which is similar to the recent *A. bispinosus* and *A. platyacanthus*; and two Plio-/Pleistocene specimens (~1.8 mya) from Italy, of which one resembles *A. irregularis pentacanthus* (Borghi 1995) and the other one is similar to *A. aranciacus* (Borghi 1995).

Fossil records are usually patchy and are likely to emerge only when a taxon becomes abundant rather than at its first appearance (Magallon 2004). For this reason, they can only mark the minimum age of a taxon, and therefore I set all fossil calibrations as minimum (lower) age constraints without an upper age limit. Also, fossil records can usually not be exactly dated but appear in geological strata of an epoch with a certain time range. In this case, I chose a conservative approach and set the lower age constraint to the younger age of the geological epoch using the geological time scale by Harland et al. (1990). Furthermore, all fossil age constraints were mapped to the appropriate stem lineage of the crown group that the fossil subtended. The crown group is considered to be the least inclusive monophyletic group that includes all extant members of the clade, whereas the stem lineage includes extinct lineages leading to the crown group that are more closely related to the crown group than to its extant sister group (Doyle and Donoghue 1993; Magallon and Sanderson 2001; Near et al. 2003).

Estimation of divergence times

Divergence times within the genus *Astropecten* were estimated using the MULTIDISTRIBUTE software package by Thorne and Kishino (2002, http://statgen.ncsu.edu/thorne/multidivtime.html). The MULTIDIVTIME program incorporated in this package uses a Bayesian approach applying Marcov chain Monte Carlo (MCMC) sampling and can accommodate multiple calibration points with upper and lower age constraints. Furthermore, it can account for evolutionary rate variation over time as well as missing data and allows for the analysis of concatenated multi-gene datasets. I followed the manual for Bayesian molecular dating using PAML/MULTIDIVTIME by Rutschmann (2005) and first estimated evolutionary model parameters for each mitochondrial DNA region separately using the program BASEML of the PAML package (Yang 1997; Yang 2007). For this step, I generated phylogenetic trees for each mtDNA region separately in MRBAYES (Ronquist and Huelsenbeck 2003) and assumed the F84 model (Kishino and Hasegawa 1989). Although this model is less parameterized than the best-fit models selected by MODELTEST, previous studies (Kishino and Hasegawa 1989, and references therein) have shown that it is actually the rate variation among sites parameter that has the greatest effect on the estimation of divergence times. I then estimated branch lengths using the program ESTBRANCHES of the MULTIDISTRIBUTE package, for which I first transformed the

BASEML output files into ESTBRANCHES input files using the program PAML2MODELINF. In a next step, I estimated branch lengths based on each set of parameters for the ML tree topology obtained earlier. All the parameters within the model, as well as the branch lengths, were estimated separately for each gene. Finally, I set upper and lower time constraints on the nodes of the tree using fossils and geological events (see above) and used the program MULTIDIVTIME to estimate divergence times and to obtain confidence intervals. Based on the oldest fossil record of *Astropecten*, *A. anglicus* (Nosowska 1997), the mean of the prior distribution for the separation time of the ingroup root from the present (rttm) was estimated to be roughly 40 Mio years. Following the manual, this value was divided by 100 Mio thus leading to an rttm of 0.4. The same value was chosen for the standard deviation (SD) of rttm (rttmsd = 0.4). An alternative value of 0.6 for rttm and rttmsd was also tried, referring to two earlier fossil records which might also belong to the genus *Astropecten*, *A. postornatus* (Rasmussen 1972) and *A.* sp (Jagt 2000), but estimated divergence times remained almost identical. Following the suggestions of the Multidivtime manual, the gamma prior distribution of the evolutionary rate of the root node (rtrate) and SD (rtratesd) was determined by consulting the branch lengths for each gene estimated by ESTBRANCHES (12S = 0.175; 16S = 0.12; COI = 0.21, over all loci = 0.17). Both rtrate and rtratesd were accordingly set equal to 0.42 substitutions at the average site per 100 Mio years. The burn-in period was 250'000 steps after which 20'000 samples were collected every 100 accepted states. For all the other parameters I chose the default options. Three independent MCMC chains were run for each calibration by using: only geological events (G), only fossil records (F), and a combined calibration with both fossils and geological events (C). The results were then compared to find possible differences in the estimated divergence times.

Results

The maximum likelihood (ML) analysis resulted in a single best tree (-lnL = 16'299.34; Figure 1) with a tree topology similar to the Bayesian and maximum parsimony (MP) trees presented in Chapter I. The ingroup was found to be monophyletic and species of *Astropecten* were grouped into three main clades: 1. species of the Indo-West Pacific; 2. species of the American east and west coast; and 3. species of the East Atlantic and Mediterranean (Figure 1). In a few cases, the relationships at nodes with low posterior probabilities (PP < 80) differed from the phylogenies in Chapter I. For instance, *A. aranciacus* grouped with the American clade in the ML analysis, although with low PP support (PP = 66), and not within the East Atlantic and Mediterranean clade as it did in the MP analysis in Chapter I. Also, even though the ML analysis grouped *A. africanus* with *A. jonstoni*, this clade was not well supported by PP (PP< 50). There were no inconsistencies between the ML and the Bayesian and MP tree topologies from Chapter I at nodes with PP and bootstrap support values above 80.

The three Marcov chain Monte Carlo (MCMC) sampling chains tended towards the same divergence times at each node, indicating sufficient sampling and burn-in periods. Divergence times estimated from G did not differ from the results of C. F resulted in divergence times of lineages 25-50% older than those inferred from G and C, whereas standard deviations (SD) of F were more than two-fold the range of those in G and C (Table 3). 95% confidence intervals (CI) in F were also more than twice the range of the CI of G and C. Having mentioned these "systematic" differences between the molecular dating in G and C versus F, I only comment on the results of G and C in the following paragraphs. Divergence times for the corresponding nodes in F can be viewed in Table 2.

Estimates of species divergence times based on the ML tree suggest that extant *Astropecten* species began to diverge in the Mid Miocene roughly 13.3 mya when Indo-West Pacific species split off from East Pacific, Atlantic and Mediterranean species (see Figure 2 and Table 2). According to my analysis, deep-sea species from the South Pacific were separated from the remaining species within the Indo-West Pacific clade (node 77) shortly afterwards ~12.2 mya, whereas the clade with the species-complex of *A. polyacanthus* (node 76) diverged ~10.8 mya. Hawaiian species, including one *A. polyacanthus* specimen from Japan, appear to have separated from the other *A. polyacanthus* specimens (node 73) ~5.2 mya.

The data further indicate that species of the American region, including *A. aranciacus*, diverged from the East Atlantic and Mediterranean species (node 103) ~11.9 mya. The divergence of *A. aranciacus* from the American species followed shortly afterwards ~10.9 mya (node 102). East Pacific and West Atlantic lineages were separated ~5.3 mya (node 98), thus somewhat before the final rise of the Isthmus of Panama ~3 mya, whereas in the West Atlantic, divergence of extant lineages already began ~7.4 mya (node 101). On the other hand, in the East Pacific, speciation of extant *Astropecten* occurred only much later (~3.3 mya), coinciding with the final rise of the Panama Isthmus.

Extant Mediterranean and East Atlantic lineages, with the exception of *A. aranciacus*, appear to have evolved in the Late Miocene ~8.4 mya, when the *irregularis–spinulosus* clade split off from the other species (node 84). While deep-sea *A. irregularis* from Portugal diverged from *A. spinulosus* and shallow-water *A. irregularis irregularis* (node 83) ~6.4 mya, *A. spinulosus* separated ~4.2 mya (node 82). *A. irregularis irregularis* and *A. irregularis pentacanthus*, two morphological variations that originally were considered to belong to the same species (Döderlein 1917; Koehler 1924; Clark and Downey 1992), diverged relatively recently (~2.0 mya; node 81). ~7.8 mya (node 80) the Mediterranean *A. jonstoni* and the Atlantic *A. africanus* split from *A. bispinosus* and *A. platyacanthus*, two exclusively Mediterranean species, and thus some time before the Messinian Salinity Crisis ~5 mya.

Discussion

The molecular phylogeny inferred from mtDNA sequence data suggests that all extant *Astropecten* lineages began to diverge in the Mid-Miocene (~13.5 mya) from a single common ancestor lineage. As the

oldest fossil record of *Astropecten* dates back to the Late Eocene (~40 million years old), this implies that other lineages whose origin predates the Mid-Miocene have gone extinct. It is, however, also possible that my sampling is incomplete, and additional old lineages are not represented in this study. I consider it unlikely though that older lineages could be revealed by more extensive sampling, because specimens from all world seas as well as the deep-sea were included, and there was a clear tendency for species to group by geographical regions. This means that missing species would probably group within the established clades and not derive from a separate older lineage. While this pattern is somewhat puzzling, results of a study on gastropods of the genus *Echinolittorina* by Williams and Reid (2004) found a similar pattern: the oldest known fossil of this group is also approximately 40 million years old (Dolin and Pacaud 2000), and a calibration of the phylogeny based on the closure of the Tethys (divergence time of Indo-West Pacific lineages) also suggests that extant lineages began to diverge from a single ancestor lineage during the Mid-Miocene. These data are in so far comparable to my results, as *Echinolittorina* also exhibits planktotrophic larvae and the genus occurs circumtropical like *Astropecten*. Although the phylogeny of *Echinolittorina* did not exhibit such a clear clustering by geographical regions as did the *Astropecten* phylogeny, the authors of this study sampled all extant species, eliminating the possibility of unsampled older lineages of extant *Echinolittorina* taxa.

In *Astropecten*, it is likely that the separation between Indo-West Pacific and East Atlantic/Mediterranen *Astropecten* lineages was caused by the closure of the Tethys Sea, with an estimated divergence time of ~13.5 mya. Thus, a fairly recent and rapid evolution has taken place within *Astropecten* and, as mentioned above, is comparable to the speciation pattern in gastropods of the genus *Echinolittorina* (Williams and Reid 2004). On the other hand, other marine invertebrate groups which have been analysed on a global level with similar molecular dating methods show much slower rates of speciation. For instance, several lineages within genera of barnacles (e.g., *Octolasmis*, *Tetraclitella* and *Verruca*; Perez-Losada et al. 2008) and also within gastropods of the genera *Nerita* (Frey and Vermeij 2008) and *Conus* (Duda 2005) appear to have diverged more than 30 mya. Moreover, in comparison to other echinoderm groups, the number of extant *Astropecten* species is remarkable for such a recent history. Species diversity is particularly striking in the East Pacific, where eight extant species diverged at the time of the final closure of the American seaway ~3 mya. Reasons for the high speciation rate within *Astropecten* are not obvious, as the presence of planktotrophic larval stages in this group would not suggest a high speciation propensity. It is, however, possible that larval dispersal is not as high as assumed, due to factors affecting the survival of the larvae and/or inconspicuous marine barriers. As for East Pacific *Astropecten*, changes in patterns of upwelling and nutrient distribution as an indirect cause of the rise of the Isthmus could have been partly responsible for the observed radiation. The onset of upwelling has also been considered as being partly responsible for radiation in strombinid gastropods in the East Pacific (Geary et al. 1992). In the West Atlantic, the chronogram suggests that factors enabling speciation must have been present already much earlier, as some West Atlantic lineages diverged several million years before the separation from the East Pacific clade. Palaeontological records of Caribbean corals and foraminifera indicate that great evolutionary changes occurred in the marine invertebrate fauna

during the Late Miocene ~6-8 mya (Keigwin 1982), which is in line with the early divergence of the *A. marginatus* and *A. cingulatus* lineages and the clade including *A. antillensis*, *A. articulaus* and *A. duplicatus* proposed by the data.

According to the inferred chronogram, the genetic separation of American *Astropecten* species on either side of the Isthmus occurred ~5.3 mya and thus some time before the final closure of the American seaway. Lineage divergence across the Panama Isthmus prior to its final rise has also been proposed for several other marine organisms (Cunningham and Collins 1994; Collins et al. 1996; Knowlton and Weigt 1998; Marko 2002). Gene flow along the region might have been disrupted earlier by shifts in currents as the Isthmus emerged. Differences in salinity between the two ocean basins were probably present before the closure ~4.7 - 4.2 mya (Keigwin 1982; Haug et al. 2001), providing potential opportunities for species adaptation.

My data suggest that gene flow between eastern and western populations of Atlantic *Astropecten* occurred until ~11 mya. Considering that the Eurasian and American continents had already drifted far apart at that point in geological history, this result is somewhat surprising. One possibility is that currents favouring dispersal between these two regions were present until the Mid-Miocene, after which they may have diminished or ceased, perhaps as a result of the contemporary closure of the Tethys. In addition, the continuously increasing distance between continental shelves might have reached a point at which gene flow across the Atlantic was no longer possible.

The results of this study further indicate that while most transatlantic lineages separated ~12 mya, *A. aranciacus*, a Mediterranean and East Atlantic species, diverged from the West Atlantic somewhat later ~11 mya. Although this could possibly suggest an ancient founder event, the relationship of this lineage is not certain, as the posterior probability (PP) of this node was very low. The weak support of this node suggests that mtDNA sequence data may not be suitable to resolve ancient phylogenetic relationships within *Astropecten*. This also applies to the deep sea *Astropecten* species of the South Pacific, which diverged around the same time and were found to group with the other Indian Ocean and Pacific species also with low nodal support. More conserved DNA regions, as can be found in nuclear DNA, would perhaps provide better phylogenetic resolution at deeper nodes and would allow for a more reliable interpretation of these ancient events. However, as several old museum specimens were used, obtaining enough nuclear DNA for amplification would have been difficult and beyond the scope of this study.

In the Mediterranean, the diversity of extant *Astropecten* species is also high and beyond that of other echinoderms. With the exception of *A. aranciacus*, my data suggest that all lineages of extant Mediterranean and East Atlantic *Astropecten* have diverged within the past 8.5 mya. There is some indication for increased radiation within the Mediterranean after the Messinian Salinity Crisis (MSC) ~5 mya, supported by the relatively recent divergence times of exclusively Mediterranean species, such as *A. bispinosus*, *A.*

platyacanthus and *A. spinulosus*, and the predominantly Mediterranean species *A. irregularis pentacanthus*. The separation of these lineages is estimated to have occurred in the Pliocene after the MSC, proposing the Mediterranean as the most likely realm of origin. Other East Atlantic and Mediterranean species, such as *A. aranciacus* and *A. irregularis* from the deep sea in Portugal, appear to have diverged before the MSC. However, their origin is not certain, since current distribution ranges favour either a Mediterranean origin with refugial populations in the Atlantic during the MSC, or an Atlantic origin with post-MSC dispersal into the Mediterranean. A study conducted on the population genetic structure in *A. aranciacus* across the Atlanto-Mediterranean region (see Chapter IV) did not reveal in which basin this species originated. Also, genetic structure in this species was found to be low, implying a high dispersal capacity even across potential marine barriers, such as the Strait of Gibraltar, allowing for rapid recolonization of either basin.

The geographic origin of the *A. jonstoni* lineage also remains unclear. This species is currently endemic to the Mediterranean but, surprisingly, seems to belong to an ancient lineage dating back ~6.9 mya. Several hypotheses can be advanced for this situation. First, *A. jonstoni* is adapted to very shallow water and potentially capable of enduring higher temperatures and salinity. As such, it could have survived through the MSC in satellite basins acting as marine refuges, as proposed by some authors to have existed in the Western Mediterranean at that time (Ben Moussa et al. 1988; Di Geronimo 1990). Others, however, consider it unlikely that any species could have survived the harsh conditions of the MSC anywhere inside the Mediterranean basin and have proposed a Messinian apocalypse (Jones 1984; Sabelli and Taviani 1984). Tethyan relics, i.e., living entities that arose during the Tethys, would therefore have had to survive outside of the Mediterranean, e.g., in the East Atlantic or Indian Ocean. This assumption leads to the second hypothesis: that *A. jonstoni* originated in the Mediterranean and managed to survive the crisis in Atlantic Ocean populations. After recolonizing the newly-suitable Mediterranean, it subsequently went extinct in the Atlantic. Although this not a very parsimonious assumption, a similar history has been suggested for other Tethyan relics, such as *Posidonia oceanica* (Taviani 2002) and some bryozoa (Rosso and Di Geronimo 1997). The last, and in my opinion most likely hypothesis, is that *A. jonstoni* originated in the Mediterranean after the MSC as species reinvaded the basin from the Atlantic. This would indicate that either the Atlantic ancestors of *A. jonstoni* went extinct after the Mediterranean colonization or that this study did not include the most closely related living species (sister species) to *A. jonstoni*. Unknown extinctions and patchy sampling are likely to obscure the origin of extant species, especially in ancient lineages where the potential for extinctions may be high.

Regarding the methods used, fossil calibration points (F) did not add further information to the geological event based calibration, as divergence time estimates were identical for the combined (C) and the purely geological (G) event based analysis. The reason for this is that G already resulted in older divergence time estimates at nodes where fossils were mapped as minimum time constraints. F, on the other hand, lead to older divergence times than G and C due to the lack of maximum time constraints. Older divergence times were also found in the study by Williams and Reid (2004) on *Echinolittorina* when using a fossil calibration

compared to a geological calibration, however, this study only used a single calibration point for each category. Contrarily, when mapping fossil records as absolute time constraints, actual divergence times are usually underestimated (Shields 2004). As the absolute ages of fossils are mostly unknown – only the geological time range of the fossil bearing rock can be determined – estimates of divergence times using fixed dating may be flawed. Additional difficulties stem from uncertainties in the identification of fossil species. Better divergence time estimates are provided if calibration points are properly constrained (Benton and Donoghue 2007). For this reason, and because high standard deviations and large ranges of 95% confidence intervals were obtained from F compared to G, I believe that estimates based on F alone have to be treated with care and are possibly less meaningful than estimates using geological events, where both upper and lower constraints were set.

Although in this study I calibrated the phylogeny using geological events that have been thoroughly investigated by the scientific community (the rise of the Isthmus of Panama and the closure of the Tethys Sea), one has to be certain that the geographical separation is the direct cause of the lineage splitting. Lineage separation may have taken place before or after the event, thus leading to possible over- or underestimation of the divergence time (Magallon 2004). My results, however, indicate that geographically separated geminates consist of several lineages, which diverged approximately at the same time. This suggests that speciation most likely took place shortly after the geographical separation.

Finally, regardless of the dating method, it is crucial that the phylogeny itself is correct when attempting to determine the origin of taxa (Near et al. 2005). As my data provided congruent tree topologies regardless of the inference method (Fig.1; cf. Chapter I), I have reason to assume that the molecular phylogeny is robust.

In summary, the calibration of phylogenies is always afflicted with inaccuracies of different sources, especially if the fossil record is not extensive and geological events can not be confidently assigned to a particular node in the tree. However, if the phylogeny shows a clear grouping of species from the same geographical region, as is the case in *Astropecten*, it is likely that these separations were caused by geological events. On the other hand, flawed identifications of fossils can often not be ruled out, especially in older and not so well preserved fossils, which are though commonly used to determine the age of a group. It is possible, for example, that since the fossil used to estimate the age of the genus *Astropecten* is not articulated, the assignment to this group is not appropriate. This would mean that the group is perhaps not as old as dating back to the Late Eocene and would explain why no extant lineage dates back past the Mid-Miocene. Further investigations though are necessary to determine if this hypothesis is valid and if a similar case could also apply to *Echinolittorina*. After all, the phylogeny of *Echinolittorina* also displays a clear separation of Indo-West Pacific lineages to the other lineages, and a plausible explanation for this could be the closure of Tethys.

Whether a fossil or a geological event based calibration, or a combination of the two, is favoured, is often guided by the author's confidence in either of the methods. Nevertheless, I believe that the tree topology can be a strong indication for whether geological events are useful for calibration, whereas fossil records have to be very well preserved and clearly assignable in order to provide a solid basis for divergence time estimates.

Acknowledgements

For providing samples and hosting museum visits I thank the following people and institutions: The late C. Ahearn and the Smithsonian National Museum of Natural History, N. Améziane and the Muséum National d'Histoire Naturelle BIMM in Paris, the Zoological Museum Amsterdam, G. Borrero and INVEMAR in Colombia, M. Brogger, A. Cabrinovich and the London Natural History Museum, T. Cerveira Borges and the BIOPESCAS team (CCMAR, University of Algarve), the California Academy of Science Collections, R. Collin, J. Collins, T. Danckwart, S. Diliberto, A. Gebruk and the Shirshov Institute, L. Granek, S. Godwin and the Bishop Museum, G. Hendler and the Natural History Museum of Los Angeles County, J. Hoover, M. Huber, M. Komatsu, D. Lane, L. Lovell and the Scripps Institution of Oceanography, C. Lüter and the Zoological Museum Berlin, H. Maag, M. Matsubara, P. Morais, C. Nielsen and the Zoological Museum of the University of Copenhagen, G. Pauley and the Florida Museum of Natural History, the Phuket Marine Biological Center, C. Ruckstuhl, H. Sekiguchi, M. Sewell, F. Solis-Marin and the Colleccion National de Equinodermos de la Universidad Nacional Autonoma de Mexico, the collections of the University of South Carolina, Q. Tahera and the Marine reference collection & Resource centre in Karachi, S. Tanner, S. Tassis, M. von Arx, K. Wakabayashi and P. Wirtz. I am grateful for the valuable advice on data analysis I received from F. Rutschmann, J. Thorne and T. Wilson, and I would also like to thank D. Blake, E. Borghi, A. Gale, H. Hess, Ch. Meier and L. Villier for assistance on paleontological issues. Further, I am greatly indebted to H. Lessios and G. Ribi who helped outlining the project and provided assistance in many ways. This project was funded in part by the Smithsonian Tropical Research Institute and the Swiss National Science Foundation.

References

Agassiz A (1877) North American starfishes. Mem Mus Comp Zool **5**: 1-136.

Ben Moussa A, Brebion P, Lauriat-Rage A, Demarcq G (1988) Interet paleobiologique des mollusques messiniens de Melilla (NE Maroc). Revue de Paléobiologie **16**: 335-358.

Benton MJ, Donoghue PCJ (2007) Paleontological evidence to date the tree of life. Mol Biol Evol **24**: 889-891. doi: 10.1093/molbev/msm017.

Blake DB (1973) Ossicle morphology of some recent asteroids and description of some west American fossil asteroids. Univ Calif Pub Geol Sci **104**, pp 59.

Blanc PL (2002) The opening of the Plio-Quaternary Gibraltar Strait: assessing the size of a cataclysm. Geodin Acta **15**: 303-317.

Borghi E (1995) Asteroidei fossili dell'Emilia (Pliocene e Pleistocene). Società Reggiana Scienze Naturali **2**: 1-8.

Cavara F (1886) Le sabbie marnose plioceniche di Mongardino e i loro fossili. Boll Soc Geol Ital **5**: 265-275.

Clark AM, Downey ME (1992) Starfishes of the Atlantic. Chapman & Hall, London, pp 794.

Coates AG, Obando JA (1996) The geologic evolution of the Central American Isthmus. In: Jackson JBC, Budd AF, G. CA (eds) Evolution and environment in tropical America. University of Chicago Press, Chicago, IL, pp 21-56.

Collins LS, Budd AF, Coates AG (1996) Earliest evolution associated with closure of the Tropical American Seaway. Proc Nat Acad Sci USA **93**: 6069-6072.

Cunningham CW, Collins TM (1994) Developing model systems for molecular biogeography: vicariance and interchange in marine invertebrates. In: Schierwater BS, Streit BB, Wagner GP, DeSalle R (eds) Molecular ecology and evolution: approaches and applications. Birkhäuser, Basel, pp 405-433.

Del Prato A (1896) Asteroidei terzari del Parmese e del Reggiano. Riv Ital Paleo Strat **2**: 42-50.

Di Geronimo I (1990) Biogeografia dello zoobenthos del Mediterraneo: origine e problematiche. Oebalia Supplement: 31-49.

Döderlein L (1917) Die Asteriden der Siboga-Expedition. I. Die Gattung *Astropecten* und ihre Stammesgeschichte. In: Brill EJ (ed) Siboga-Expeditie. Uitkomsten op zoölogisch, botanisch, ozeanographisch en geologisch gebied verzameld in Nederlandsch Oost-Indie 1899-1900 aan boord H.M. "Siboga". 46 (a), Leiden, pp 191.

Doyle JA, Donoghue MJ (1993) Phylogenies and angiosperm diversification. Paleobiology **19**: 141-167.

Felsenstein J (1988) Phylogenies from Molecular Sequences - Inference and Reliability. Ann Rev Genet **22**: 521-565.

Fisher WK (1906) New Starfishes from the Pacific Coast of North America. Proc Wash Acad Sci **8**: 11-139.

Fisher WK (1911) Asteroidea of the North Pacific and adjacent waters. Part 1. Phanerozonia and Spinulosa. Bull U.S. Nat Mus **76**, pp 420.

Fisher WK (1913) Four new genera and fifty-eight new species of starfishes from the Philippine Islands, Celebes, and the Moluccas. Proc U.S. Nat Mus **43**: 599-648.

Frey MA, Vermeij GJ (2008) Molecular phylogenies and historical biogeography of a circumtropical group of gastropods (Genus: *Nerita*): Implications for regional diversity patterns in the marine tropics. Mol Phylogenet Evol **48**: 1067-1086. doi: 10.1016/j.ympev.2008.05.009.

Geary DH, Brieske TA, Bemis BE (1992) The influence and interaction of temperature, salinity and upwelling of the stable isotopic profiles of strombid gastropod shells. Palaios **7**: 77-85.

Gray JE (1841) A synopsis of the genera and species of the class Hypostoma (Asterias, Linnaeus). Ann Mag Nat Hist **1**: 175-184.

Harland WB, Armstrong RL, Cox AV, Craig LE, Smith AG, Smith DG (1990) A geologic time scale, 1989 edition. University Press, Cambridge, pp 293.

Haug GH, Tiedemann R, Zahn R, Ravelo AC (2001) Role of Panama uplift on oceanic freshwater balance. Geology **29**: 207-210.

Heller C (1858) Über neue fossile Stelleriden. Sitzungsberichte der Kaiserlichen Akademie der Wissenschaften. Mathematisch-naturwissenschaftliche Classe, Abteilung I **28**: 155-170.

Hofrichter R (2003) Das Mittelmeer. Fauna, Flora, Ökologie. Band I: Allgemeiner Teil. Spektrum Akademischer Verlag, Heidelberg - Berlin, pp 607.

Jagt WM (2000) Late Cretaceous-Early Palaeogene echinoderms and the K/T boundary in the southeast Netherlands and northeast Belgium. Part 5: Asteroids. Scripta Geologica **121**: 377-503.

Jones CC (1984) Messinian refugia: evidence of some Venerinae bivalves. Interim colloquium on Mediterranean Neogene marine megafaunal palaeoenvironments and biostratigraphy. Ann géolog Pays Hellen **17**: 69-77.

Kaczmarska G (1987) Asteroids from the Korytnica Basin (Middel Miocene; Holy Cross Mountains, Central Poland). Acta Geologica Polonica **37**: 131-144.

Keigwin L (1982) Isotopic paleo-oceanography of the Caribbean and East Pacific - role of the Panama uplift in Late Neogene time. Science **217**: 350-352.

Kishino H, Hasegawa M (1989) Evaluation of the maximum-likelihood estimate of the evolutionary tree topologies from DNA-sequence data, and the branching order in homonoidea J Mol Evol **29**: 170-179.

Knowlton N, Weigt LA (1998) New dates and new rates for divergence across the Isthmus of Panama. Proc R Soc Lond Ser B-Biol Sci **265**: 2257-2263.

Knowlton N, Weigt LA, Solorzano LA, Mills DK, Bermingham E (1993) Divergence in proteins, mitochondrial DNA, and reproductive compatibility across the Isthmus of Panama. Science 260: 1629-1632.

Koehler R (1909) Echinodermes provenant des campagnes du yacht Princesse-Alice. Resultats des campagnes scientifiques (Monaco) **34**: 1-317.

Koehler R (1910) Shallow-Water Asteroidea. Echinoderma of the Indian Museum, Calcutta, Indian Museum **191**, pp206.

Koehler R (1924) Les Echinodermes des Mers D'Europe. Librarie Octave Doin, Paris, pp 210.

Krijgsman W, Hilgen FJ, Raffi I, Sierro FJ, Wilson DS (1999) Chronology, causes and progression of the Messinian salinity crisis. Nature **400**: 652-655.

Kroh A, Harzhauser (1999) An echinoderm funa from the Lower Miocene of Austria: paleoecology and implications for Central Paratethys paleobiogeography. Ann Nat hist Mus Wien **101A**: 145-191.

Kudrin LN (1957) Palaeoecological investigations of the lowermost Tortonian depostis from the south-western part of the Russian Platform [in Russian]. Geologiceskij Sbornik **4**: 277-294.

Ludwig H (1897) Die Seesterne des Mittelmeeres. Fauna und Flora des Golfes von Neapel und der angrenzenden Meeres-Abschnitte. Zool Stn Neapel **24**, pp 491.

Magallon S, Sanderson MJ (2001) Absolute diversification rates in angiosperm clades. Evolution **55**: 1762-1780.

Magallon SA (2004) Dating lineages: Molecular and paleontological approaches to the temporal framework of clades. Int J Plant Sci **165**: S7-S21.

Marko PB (2002) Fossil calibration of molecular clocks and the divergence times of geminate species pairs separated by the Isthmus of Panama. Mol Biol Evol **19**: 2005-2021.

Müller J, Troschel FH (1842) System der Asteriden. Bieweg & Sohn, Braunschweig, pp 132.

Near TJ, Kassler TW, Koppelman JB, Dillman CB, Philipp DP (2003) Speciation in North American black basses, *Micropterus* (Actinopterygii : Centrarchidae). Evolution **57**: 1610-1621.

Near TJ, Meylan PA, Shaffer HB (2005) Assessing concordance of fossil calibration points in molecular clock studies: An example using turtles. Am Nat **165**: 137-146.

Nosowska E (1997) Asteroids from the Nawadzice Sands (Middle Miocene; Holy Cross Mountains, Central Poland). Acta Geologica Polonica **47**: 225-241.

Pereira P, Bajo I, Aguirre J (2002) Primeros datos sobre la presencia de *Astropecten* (Echinodermata: Asteroidea) en el Mioceno de Cerro Palomar (Espera, Càdiz). XVIII Jornadas de la Soc. Esp. de Paleo., Salamanca, Libro de resumenes: 200-201.

Pereira P, Cachão M, Silva CM (2003) Astereoida (Echinodermata) do Miocénico da Baixo Tejo-Sado. Ciências da Terra, n.° esp. **5**: A106-A109.

Perez-Losada M, Harp M, Hoeg JT, Achituv Y, Jones D, Watanabe H, Crandall KA (2008) The tempo and mode of barnacle evolution. Mol Phylogenet Evol **46**: 328-346. doi: 10.1016/j.ympev.2007.10.004.

Perez-Losada M, Hoeg JT, Crandall KA (2004) Unraveling the evolutionary radiation of the thoracican barnacles using molecular and morphological evidence: A comparison of several divergence time estimation approaches. Syst Biol **53**: 244-264. doi: 10.1080/10635150490423458.

Perrier E (1875/6) Révision de la collection de Stellerides du Muséum d'Histoire Naturelle de Paris. Archives de Zoologie Experimentale et Génerale 4: 265-450.

Posada D, Crandall KA (1998) MODELTEST: testing the model of DNA substitution. Bioinformatics **14**: 817-818.

Rasmussen HW (1972) Lower Tertiary Crinoidea, Asteroidea and Ophiuridea from Northern Europe and Greenland. Biologiske Skrifter **19**: 1-83.

Rico-García A, Bajo I, Pereira P (2008) El Genero *Astropecten* (Echinodermata, Asteroidea) en el Neogeno superior del oeste de la Cuenca del Guadelquivir. Studia Geologica Salamanticensia. Volumen especial **8**: 53-69.

Rögl F (1999) Mediterranean and Paratethys. Facts and Hypothesis of an Oligocene to Miocene Paleogeography (short overview). Geologica Carpathica **50**: 339-349.

Rögl F, Steininger F (1983) Vom Zerfall der Tethys zu Mediterran und Paratethys. Ann Nat hist Mus Wien **85A**: 135-163.

Ronquist F, Huelsenbeck JP (2003) MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics **19**: 1572-1574.

Rosso A, Di Geronimo I (1997) Deep-sea Pleistocene Bryozoa of Southern Italy. Geobios **30**: 303-317.

Rutschmann F (2005) Bayesian moleuclar dating using PAML/MULTIDIVTIME. A step-by-step manual. Version 1.5 (July 2005). Available at http://www.plant.ch.

Sabelli B, Taviani M (1984) The paleobiogeographic distribution of the Mediterranean benthic mollusks and the Messinian Salinity Crisis or where did the mollusks go? Interim Colloquium on Mediterranean Neogene Marine Megafaunal Paleoenvironments and Biostratigraphy. Ann géolog Pays Hellen **17**: 263-269.

Sacco F (1893) Sopra alcuni Asteroidei fossili. Atti R Accad Sci Torino **28** (1892-93): 407-445.

Say T (1825) On the species of the Linnean genus Asterias inhabiting the coast of the United States. J Acad Nat Sci Phil **5**: 151-154.

Shields R (2004) Pushing the envelope on molecular dating. Trends Genet **20**: 221-222. doi: 10.1016/j.tig.2004.03.011.

Siesser WG (1980) Late Miocene origin of the Benguela upwelling system off Northern Namibia. Science **208**: 283-285.

Sladen WP (1889) Report on the Asteroidea. Report on the scientific results of the voyage of the H. M. S. Challenger during the years 1873-1876. Zoology **30**: 1-893.

Smith AB, Peterson KJ (2002) Dating the time of origin of major clades: Molecular clocks and the fossil record. Ann Rev Earth Planet Sci **30**: 65-88.

Smith AB, Pisani D, Mackenzie-Dodds JA, Stockley B, Webster BL, Littlewood TJ (2006) Testing the molecular clock: Molecular and paleontological estimates of divergence times in the echinoidea (Echinodermata). Mol Biol Evol **23**: 1832-1851. doi: 10.1093/molbev/msl039.

Swofford DL (2003) PAUP*. Phylogenetic Analysis Using Parsimony (*and other Methods). Version 4, Sunderland, Massachusetts.

Taviani M (2002) The Mediterranean benthos from late Miocene up to present: ten million years of dramatic climatic and geologic vicissitudes. Biol Mar Medit **9**: 445-463.

Thorne JL, Kishino H (2002) Divergence time and evolutionary rate estimation with multilocus data. Syst Biol **51**: 689-702. doi: 10.1080/10635150290102456.

Tinn O, Oakley TH (2008) Erratic rates of molecular evolution and incongruence of fossil and molecular divergence time estimates in Ostracoda (Crustacea). Mol Phylogenet Evol **48**: 157-167. doi: 10.1016/j.ympev.2008.03.001.

Verrill AE (1914) Monograph of the shallow-water Starfishes of the North Pacific Coast from the Arctic Ocean to California. Harriman Alaska Series **14**: 1-420.

Verrill AE (1915) Report on the Starfishes of the West Indies, Florida, and Brazil. Bull Lab Nat Hist **7**: 2-232.

White BN (1994) Vicariance biogeography of the open-ocean Pacific. Prog Oceanogr **34**: 257-284.

Wilke T, Rolan E, Davis GM (2000) The mudsnail genus *Hydrobia* s.s. in the northern Atlantic and western Mediterranean: a phylogenetic hypothesis. Mar Biol **137**: 827-833.

Williams ST, Reid DG (2004) Speciation and diversity on tropical rocky shores: A global phylogeny of snails of the genus Echinolittorina. Evolution **58**: 2227-2251.

Wood AR, Apte S, MacAvoy ES, Gardner JPA (2007) A molecular phylogeny of the marine mussel genus Perna (Bivalvia : Mytilidae) based on nuclear (ITS1&2) and mitochondrial (COI) DNA sequences. Mol Phylogenet Evol **44**: 685-698. doi: 10.1016/j.ympev.2006.12.019.

Yang ZH (1997) PAML: a program package for phylogenetic analysis by maximum likelihood. Comput Appl Biosci **13**: 555-556.

Yang ZH (2007) PAML 4: Phylogenetic analysis by maximum likelihood. Mol Biol Evol **24**: 1586-1591. doi: 10.1093/molbev/msm088.

Yang ZH, Yoder AD (2003) Comparison of likelihood and Bayesian methods for estimating divergence times using multiple gene loci and calibration points, with application to a radiation of cute-looking mouse lemur species. Syst Biol 52: 705-716. doi: 10.1080/10635150390235557.

Zuckerkandl E, Pauling L (1965) Molecules as documents of evolutionary history. J Theor Biol **8**: 357-366.

Table 1: Species, identified code, location and date of sampling, voucher and GenBank accession codes for *Astropecten* and outgroup specimens.

Species	ID	Location/Date	Voucher	Genbank accession codes		
				12S rRNA	16S rRNA	COI
A. africanus	Aafri - STo1	EA - São Tomé/ Feb 2006		FJ171765	FJ177591	FJ195695
A. alligator	Aalli - Col1	WA - Santa Maria, Colombia; 300 m/ 2001	INV. EQU01809	FJ171787	FJ177548	
A. americanus	Aamer - Flo1	WA - Tampa Bay, Florida, USA; 271 m/ Mar 2003	UF 3471	FJ171786	FJ177547	
A. antillensis	Aanti - Pan1	WA - San Blas, Panama/ Feb 2002		FJ171794	FJ177541	FJ195654
A. aranciacus	Aaran - Far3	EA - Faro, Faro Portugal; 30 m/ Jun 2005		FJ171770	FJ177596	FJ195669
A. armatus	Aarma - Mex1	EP - Puerto Peñasco, Northern Sea of Cortez, Mexico; 69 m/ Mar 1985	UNAM 4190	FJ171785	FJ177563	
A. articulatus	Aarti - Pan3	WA - Bocas del Toro, Panama/ 1996		FJ171792	FJ177544	FJ195660
	Aarti - SCa1	WA - Cape Island, South Carolina, USA; 12-13 m	USC S713	FJ171795	FJ177545	
A. bispinosus	Abisp - Sar1	WM - Costa Colostrai, Sardinia; 17 m/ Aug 2002		FJ171745	FJ177570	FJ195733
A. cingulatus	Acing - Pan1	WA - Isla Escuda de Veraguas, Panama; 42-39 m/ Aug 2004		FJ171802	FJ177552	FJ195664
A. comptus	Acomp - Flo1	WA - Gulf of Mexico, off St. Petersburg, Florida; 116 m/ Nov 2004	UF 3249	FJ171788	FJ177551	
A. duplicatus	Adupl - Bim1	WA - Bimini, Bahamas; 0 m/ Feb 2003		FJ171797	FJ177537	FJ195717
A. erinaceus	Aerin - Pan1	EP - Isla Montuosa, Golf of Chiriqui, Panama; 42.6 m/ May 2004		FJ171778	FJ177556	FJ195656
A. granulatus	Agran - Aus1	IP - Cobourg Peninsula, Northern Territory, Australia; 13 m/ Sep 1985	SI NMNH E38949	FJ171825	FJ177620	FJ195685
A. indicus	Aindi - Pak2	IO - Clifton, Karachi, Pakistan/ 2005		FJ171820	FJ177617	FJ195691
	Aindi - Bru1	IP - Serasa, Brunei/ Dec 2004		FJ171821	FJ177618	FJ195726
	Aindi - Tha1	IP - Phuket, Thailand/ Apr 1997	PMBC 19223, *A. monacanthus*	FJ171819	FJ177619	FJ195692
A. irregularis irregularis	Airre - Sar3	WM - Costa Colostrai, Sardinia; 0-30m/ Aug 2002		FJ171750	FJ177575	FJ195678
A. irregularis	Airre - Far1	EA - off Faro, Portugal; 530-540 m/ May 2005		FJ171758	FJ177583	FJ195665
A. irregularis pentacanthus	Airre - Kav2	EM - Kavala, Greece; > 50 m/ Mar 2006		FJ171753	FJ177578	FJ195697
A. javanicus	Ajava - Bru1	IP - Serasa, Brunei/ Dec 2004		FJ171824	FJ177616	FJ195725
A. jonstoni	Ajons - Sar1	WM - Costa Colostrai, Sardinia/ Aug 2000		FJ171762	FJ177588	FJ195703
A. latespinosus	Alate - Jap1	WP - Toyama Bay, Japan/ 2004		FJ171829	FJ177623	FJ195722
A. marginatus	Amarg - Col1	WA - Santa Marta, Colombia/ 2003	INV EQU02562	FJ171805	FJ177555	FJ195715
A. monacanthus	Amona - Aus1	IP - Torres Strait, Queensland, Australia; 0 m/ Jun 1979	SI NMNH E35296	FJ171826	FJ177621	FJ195735
A. nitidus	Aniti - Col1	WA - Santa Marta, Colombia; 153 m/ 2001	INV EQU01841	FJ171789	FJ177550	FJ195716
A. oerstedii	Aoers - Pan1	WA - San Blas, Panama/ Feb 2002		FJ171782	FJ177561	FJ195655
A. platyacanthus	Aplat - Cyp2	EM - De Capo Bay, Cyprus; 3 m/ Oct 2004		FJ171748	FJ177573	FJ195724
A. polyacanthus	Apoly - Mau2	CP - Kanaio, Maui, Hawaii; 10 m/ Feb 2006		FJ171808	FJ177602	FJ195694
	Apoly - Sey1	IO - Picard Island, Aldabra Islands, Seychelles; reef flat/ Mar 1987	SI NMNH E35107	FJ171810	FJ177607	FJ195686
	Apoly - Fij1	SP - Bligh Water, Fiji; 143-173 m/ Aug 1998	MNHN EcAh 4734	FJ171827		
	Apoly - NZ1	SP - Auckland, New Zealand/ 2003		FJ171813	FJ177610	FJ195711
	Apoly - Jap1	WP - Toyama Bay, Japan/ Nov 2003		FJ171806	FJ177605	FJ195719
	Apoly - Phi1	IP - Aligbay, Mindanao, Philippines; 1.5 m/ May 1979	SI NMNH E48900	FJ171812	FJ177609	FJ195736
	Apoly - Dub1	IO - Dugass Beach, Dubai, U.E.A./ Feb 1981	SI NMNH E35065	FJ195652	FJ177633	
	Apoly - Mar1	SP - Ua Pou, Marquesas Islands; 70-77 m/ Aug 1997	MNHN EcAh 4748	FJ969165		
A. regalis	Arega - CRi1	EA - Coco, Guanacaste, Costa Rica; 4 m/ 1933	LACM 1933-123	FJ195652	FJ177633	
A. scoparius	Ascop - Jap1	WP - Toyama Bay, Japan/ Nov 2003		FJ171818	FJ177613	FJ195718
A. siderealis	Averr - Pan1	EP - Coiba Island, Panama; 58 m/ May 2004		FJ171779	FJ177557	FJ195661
A. sp. 1	Asp1 - Fij1	SP - Malolo, Viti Levu, Fiji; 39 m/ Oct 1998	MNHN EcAh 4730	FJ171817		
A. sp. 2	Asp2 - Fij1	SP - SE Viti Levu, Fiji; 244-252 m/ Aug 1998	MNHN EcAh 4725	FJ171836	FJ177627	FJ195702
A. sp. 3	Asp3 - Fij2	SP - off Suva, Fiji; 478-500m/ Mar 1999	MNHN EcAh 4735	FJ171832	FJ177625	FJ195705
A. sp. 4; (possibly *A. tasmanicus* or *A. eremicus*)	Asp4 - Fij1	SP - S Namenalala, Fiji; 364-369 m/ Mar 1999	MNHN EcAh 4738	FJ171833	FJ177626	FJ195707
A. sp. 5	Asp5 - Ton1	SP - SW Tongatapu, Tonga; 319-333 m/ Jun 2000	MNHN EcAh 4741	FJ171835		FJ195708
A. sp. 6	Asp6 - Phi1	IP - Leyte Island, Philippines; 76 m/ Nov 1979	SI NMNH E53739	FJ171828	FJ177622	FJ195737
A. sp. 7	Asp7 - Aus1	SP - Pallarenda Beach, Townsville, Australia; 0 m/ May 2004		FJ171823	FJ177615	FJ195728
A. spinulosus	Aspin - Cre1	EM - Cres, Croatia; 5 m/ Oct 2002		FJ171756	FJ177581	FJ195706
A. triseriatus	Atris - Oah1	EP - Kailua, Oahu, Hawaiian Islands; 22m/ 1980	BM 1980.536	FJ171809	FJ177604	
A. vappa	Avapp - Bru1	IP - Serasa, Brunei/ Dec 2004		FJ171816	FJ177612	FJ195727
A. verilli	Averr - Pan1	EP - Panama/ 1996?		FJ171779	FJ177557	FJ195661
	Averr - Cal1	EP - Monterrey Bay, California, USA/ May 1996	Calacad 105628	FJ171783	FJ177560	FJ195713
	Averr - Cal2	EP - Point Loma, San Diego, USA; 220 m/ Nov 2002	SIO BIC E3481	FJ171781	FJ177559	FJ195730
A. zebra	Azebr - PNG1	SP - Deboin Mission, SE Papua New Guinea,/ Jun 1979	SI NMNH E50681	FJ171822	FJ177614	FJ195689
Tethyaster sp. 1	Astro - Ton1	SP - Eua, Tonga; 463-464m/ Jun 2000	MNHN EcAh 4758	FJ171840		FJ195709
Tethyaster sp. 2	Astro - Naz1	EP - Nazca submarine ridge; 230-280 m/ May 1987		FJ171841	FJ177630	FJ195734

Ctenophoraster sp.	Cphor - Mar1	SP - Fatu Hiva, Marquesas Islands; 85-130m/ Sep 1997	MNHN EcAh 4749	FJ171837		
Ctenopleura sp.	Cpleu - Fij1	SP - Bligh Water, N Viti Levu, Fiji; 143-173 m/ August 1998	MNHN EcAh 4732	FJ171843	FJ177632	
Pseudarchaster parelii	Ppare - Pan1	WA - San Blas, Panama/ Feb 2003		FJ171838	FJ177628	FJ195738
Tethyaster subinermis	Tsubi - Far1	EA - off Portimão, Faro, Portugal; 120-131m/ May 2005		FJ171839	FJ177629	FJ195672
Thrissacanthias penicillatas	Tpeni - Die1	EP - off San Diego, California, USA; 1215 m/ Oct 2005	SIO BIC E3857	FJ171842	FJ177631	FJ195673
Total	**60**					

EA = East Atlantic; WM = West Mediterranean; EM = East Mediterranean; WA = West Atlantic; EP = East Pacific; CP = Central Pacific; SP = South Pacific; WP = West Pacific; IP = Indo-Pacific; IO = Indian Ocean.
BM = Bishop Museum, Honolulu, Hawaii; Calacad = California Academy of Science; INV = INVEMAR, Instituto de Investigaciones Marinas y Costeras, Colombia;
LACM = Los Angeles County Museum; MNHN = Musée nationale d'histoire naturelle, Paris, Echinoderm Collection; PMNH = Phuket Museum of Natural History;
SI NMNH = Smithsonian National Museum of Natural History, Invertebrate Collection; SIO BIC = Scripps Institution of Oceanography - Benthic Invertebrate Collection; UF = Florida Museum of Natural History; UNAM = Universidad Nacional Autonoma de Mexico, Colleccion Nacional de Equinodermos; USC = University of South Carolina.

Table 2: Geological events (G) and fossil records (F) used to set time constraints on nodes of the inferred molecular phylogeny of *Astropecten.*

Geological event	**Reference**	**Time Period/Locality**	**Node**	**Age mya**
Final rise of the Panama Isthmus	Coates 1996	Pliocene/Panama	98	L 2.8 U 6
Closure of the Tethys	Rögl 1983	Mid to late Miocene/ eastern Mediterranean	104	L 11.8 U 18
Fossil record				
A. anglicus	Nosowska 1997	Late Eocene, Bartonian/England	root	40
A. bispinosus/ *A. platyacanthus*	Sacco 1893	Pliocene/ Italy	80	L 1.8
A. irregularis pentacanthus	Borghi 1995	lower Pleistocene, upper Pliocene/ Italy	82	L 1.8
A. aranciacus	Borghi 1995	lower Pleistocene, upper Pliocene/ Italy	103	L 1.8

Table 3: Actual ages of nodes, standard deviations (SD) and 95% confidence intervals (5%, 95% CI) estimated using Multidivtime for a molecular phylogeny of *Astropecten* species with a calibration using geological events (G) and a calibration using fossil records (F). Nodes 0-52 correspond to the terminal nodes of specimens sampled at present time (not indicated in Figure 2). Location of nodes 53-104 can be viewed in Figure 2.

	Geological events (G)			Fossil records (F)		
Node	**Actual time**	**SD**	**5%, 95% CI**	**Actual time**	**SD**	**5%, 95% CI**
0-52	0	0	0, 0	0	0	0, 0
53	3.70	1.01	1.97, 5.96	4.79	2.76	1.65, 11.78
54	7.09	1.34	4.72, 9.99	9.10	4.86	3.52, 21.55
55	10.95	1.42	8.50, 14.13	13.97	7.25	5.67, 32.54
56	1.84	0.50	0.96, 2.95	2.40	1.38	0.82, 5.96
57	2.22	0.57	1.22, 3.46	2.88	1.62	1.02, 7.04
58	1.98	0.57	1.00, 3.25	2.67	1.63	0.87, 6.89
59	2.60	0.77	1.39, 4.39	3.57	2.32	1.14, 9.65
60	6.96	1.34	4.54, 9.78	9.13	4.98	3.43, 21.86
61	9.20	1.28	6.95, 12.06	11.85	6.18	4.79, 27.47
62	2.14	1.21	0.21, 4.96	2.88	2.41	0.25, 9.14
63	1.00	0.98	0.05, 4.01	1.76	2.39	0.07, 8.59
64	6.36	1.34	3.98, 9.25	8.14	4.56	3.00, 19.85
65	9.66	1.30	7.41, 12.58	12.42	6.46	5.06, 28.65
66	2.75	1.50	0.37, 6.15	3.58	2.81	0.35, 10.80
67	4.49	0.98	2.80, 6.66	5.76	3.18	2.15, 13.78
68	0.74	0.48	0.04, 1.81	0.96	0.83	0.05, 3.08
69	1.57	0.49	0.75, 2.64	2.01	1.19	0.63, 5.09
70	2.15	0.69	0.86, 3.60	2.78	1.69	0.77, 7.09
71	2.92	0.67	1.78, 4.42	3.78	2.11	1.37, 9.11
72	4.21	0.93	2.60, 6.23	5.44	3.00	2.00, 13.07
73	5.16	0.96	3.52, 7.31	6.71	3.64	2.61, 16.01
74	6.91	1.10	5.02, 9.34	8.91	4.73	3.58, 20.89
75	8.26	1.22	6.19, 10.98	10.58	5.53	4.29, 24.59
76	10.82	1.32	8.60, 13.83	13.89	7.16	5.72, 32.23
77	12.16	1.33	10.02, 15.32	15.54	7.97	6.43, 35.69
78	3.73	0.78	2.37, 5.44	4.89	2.73	1.84, 11.80
79	6.91	1.21	4.66, 9.47	8.99	4.83	3.47, 20.98
80	7.84	1.14	5.82, 10.31	10.20	5.35	4.15, 23.55
81	2.01	0.53	1.13, 3.18	2.61	1.51	0.92, 6.48
82	4.25	0.84	2.79, 6.11	5.53	2.99	2.17, 13.11
83	6.41	1.01	4.67, 8.63	8.35	4.44	3.34, 19.44
84	8.50	1.16	6.49, 11.05	11.04	5.75	4.52, 25.55
85	0.88	0.29	0.38, 1.53	1.20	0.76	0.35, 3.15
86	2.62	0.57	1.60, 3.82	3.58	2.00	1.30, 8.65
87	3.61	0.78	2.26, 5.33	4.96	2.77	1.78, 12.03
88	1.35	0.67	0.20, 2.79	1.94	1.51	0.23, 5.79
89	0.44	0.32	0.02, 1.22	0.64	0.65	0.02, 2.33
90	4.20	0.72	2.74, 5.53	6.09	3.43	2.22, 14.85
91	0.69	0.38	0.07, 1.50	0.97	0.80	0.07, 3.05
92	0.65	0.24	0.22, 1.17	0.91	0.60	0.22, 2.45
93	1.56	0.37	0.92, 2.35	2.19	1.29	0.76, 5.46
94	1.94	0.47	1.09, 2.95	2.74	1.61	0.93, 6.81
95	2.68	0.51	1.74, 3.74	3.78	2.13	1.40, 9.18
96	2.98	0.52	2.03, 4.06	4.21	2.34	1.60, 10.15
97	3.29	0.54	2.28, 4.42	4.64	2.56	1.78, 11.17
98	5.35	0.49	4.20, 5.98	7.82	4.19	3.07, 18.44
99	6.21	0.68	4.87, 7.58	8.74	4.62	3.49, 20.60
100	6.89	0.77	5.46, 8.52	9.54	5.03	3.83, 22.21
101	7.40	0.85	5.85, 9.19	10.14	5.32	4.07, 23.75
102	10.93	1.19	8.91, 13.67	14.33	7.38	5.93, 33.18
103	11.87	1.23	9.86, 14.73	15.41	7.89	6.40, 35.44
104	13.32	1.24	11.85, 16.46	17.01	8.68	7.07, 38.85

Figure 1: Maximum likelihood tree of *Astropecten* rooted by the outgroup specimens using the GTR+I+Γ evolutionary model (base frequencies: A=0.32, C=0.20, G=0.13, T=0.34; rate matrix: A-C=1.19, A-G=9.16, A-T=1.27, C-G=0.37, C-T=9.22, G-T=1; proportion of invariable sites I=0.52; gamma shape Γ= 0.78); Numbers on branches indicate node credibility and refer to posterior probabilities inferred from Bayesian analysis.

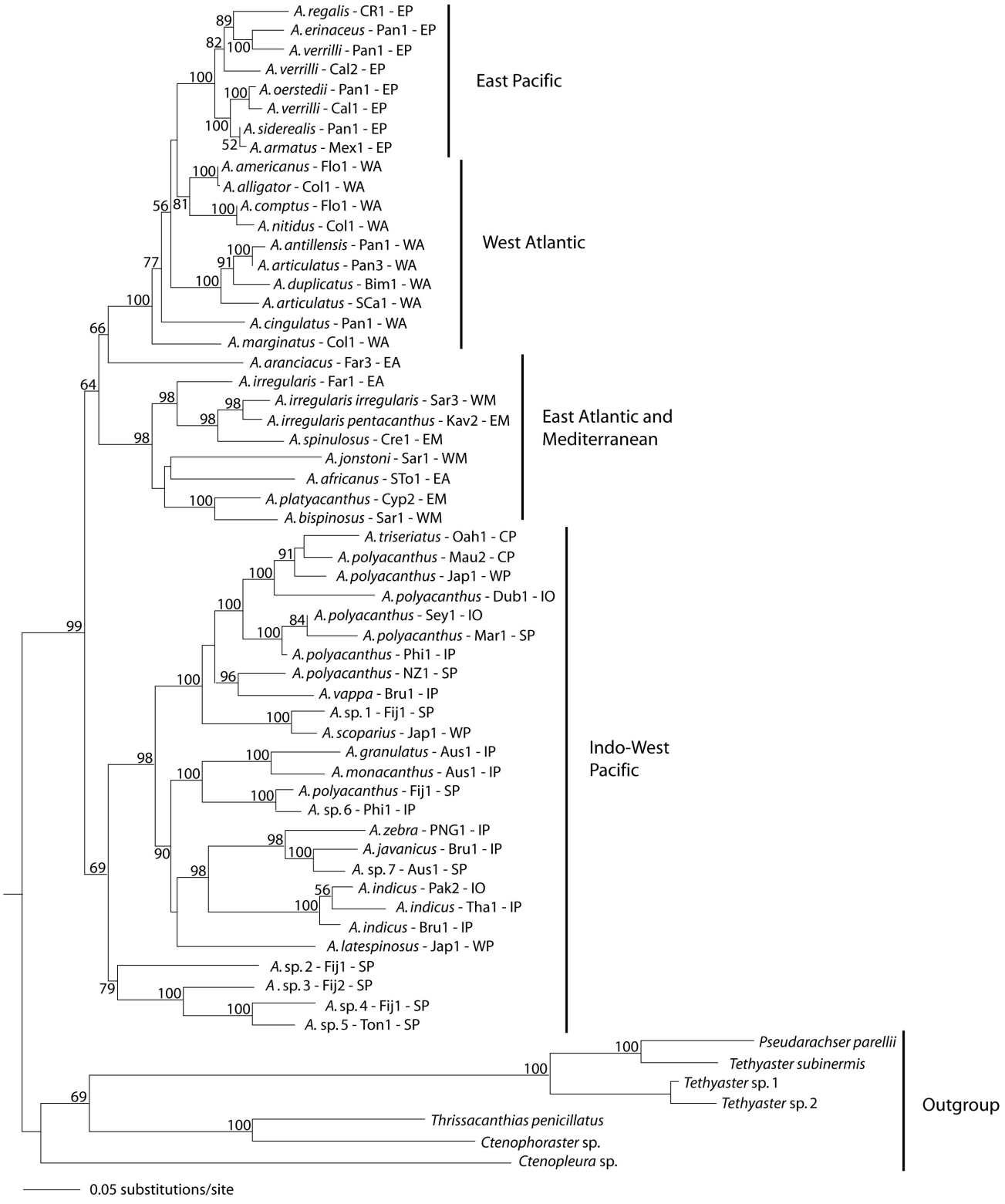
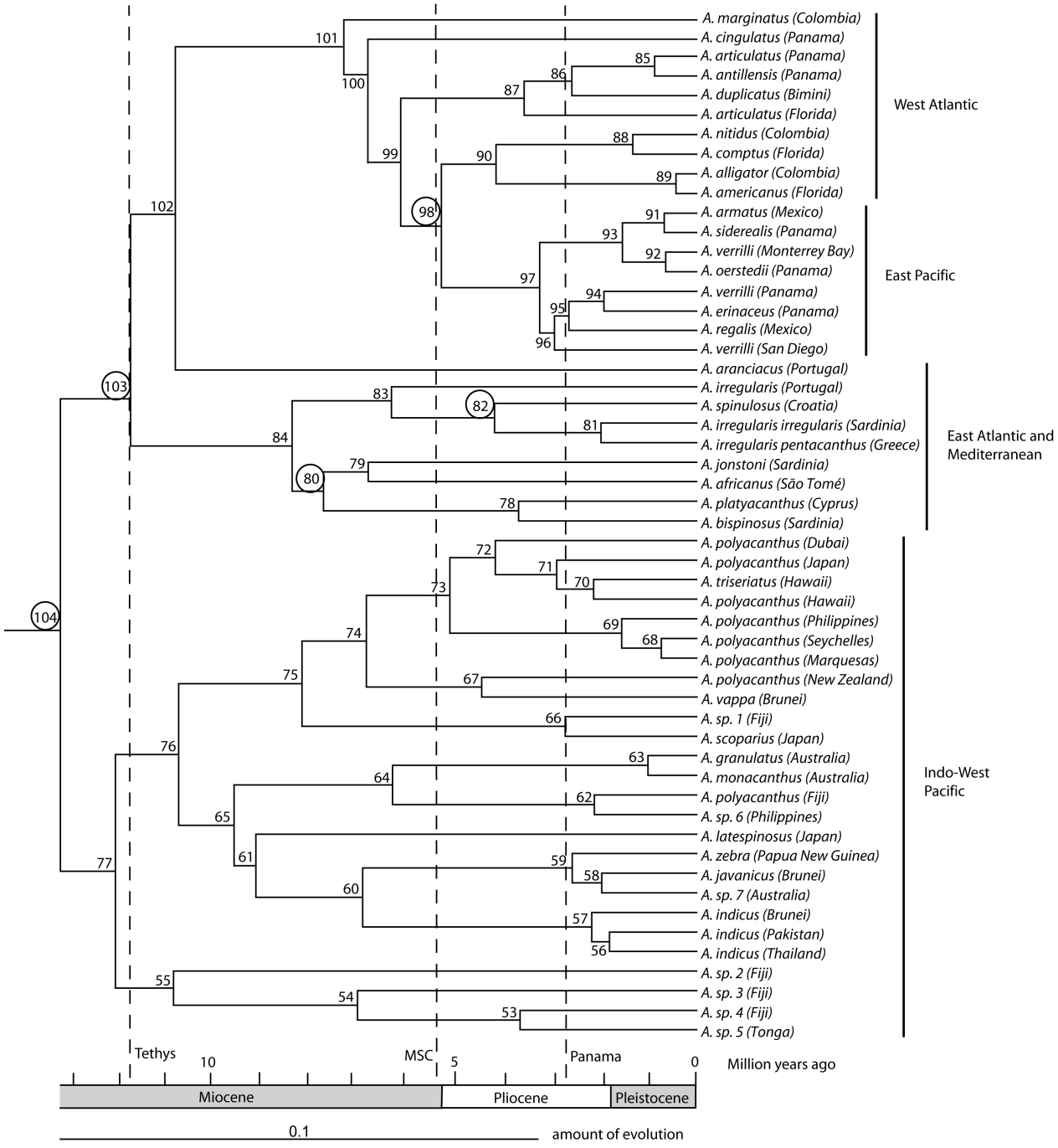


Figure 2: Chronogram of *Astropecten* calculated using MULTIDIVTIME based on a calibration with geological events (G). Outgroup specimens are not shown for illustration purposes. Terminal nodes are labelled with numbers 0 - 52 (not shown) and are detailed in Table 3; internal nodes are indicated in the chronogram with numbers 53 - 104. Nodes where calibrations were set are highlighted by a circle. Dashed lines indicate the final closure of the Tethys, the end of the Messinian Salinity Crisis (MSC) and the final rise of the Panama Isthmus.



CHAPTER III

Characterization of nine microsatellite loci in the sea star *Astropecten aranciacus* and cross-species amplification for related taxa

Deborah E. Zulliger, Markus Ruch, Samuel Tanner and Georg Ribi

Abstract

So far, only few microsatellite markers have been developed and extensively tested for echinoderms. To study the population genetic structure of the sea star *Astropecten aranciacus,* we developed primers for nine polymorphic microsatellite loci and tested them on two populations from Faro in Portugal (N = 25) and from La Herradura in Spain (N = 20). Within populations, allele numbers varied from four to 20, while expected and observed heterozygosities ranged from 0.593 to 0.936 and from 0.222 to 0.900, respectively. Additional cross-species amplifications were polymorphic at some loci, indicating their potential usefulness for related taxa.

The sea star *Astropecten aranciacus* was once abundant in the Mediterranean Sea (Burla *et al*. 1972). Because of its large body size of up to 60 cm in diameter and its formerly high population densities, it is believed to be an important predator of the benthos. Recently, however, a decline in populations of *A. aranciacus* has been observed in several areas within the Mediterranean Sea (G. Ribi, unpublished; H. Lessios, H. Massé, H. Moosleitner, L. Santella, personal communication). The reasons for this decline are not yet fully understood, but intense fishing (trawling, gill nets), pollution and increased water temperatures presumably facilitating wasting disease in echinoderms (G. Ribi *et al*. unpublished) are the most obvious explanations. To estimate the genetic variability and population structure of *A. aranciacus,* we developed nine microsatellite loci. Development of suitable microsatellite loci for echinoderms and particularly for asteroids has not been very extensive so far. This may be related to difficulties in microsatellite isolation similar to the ones encountered in other groups of invertebrates. To our knowledge, the only microsatellite loci characterized for asteroids are for *Acanthaster planci* (Yasuda *et al*. 2006). In this note, we characterize nine microsatellite markers in the sea star *A. aranciacus* and show their successful amplification in related taxa.

Genomic DNA was first extracted from approximately 40 mg of a 96% ethanol-preserved arm tip containing tube feet of an individual using a DNeasy Tissue Kit (QIAGEN). An enriched microsatellite library was developed by ecogenics GmbH. In brief, fragments were size selected following enzymatic digestion, ligated into TSPAD-linker (Tenzer *et al*. 1999) and enriched by magnetic bead selection with biotin-labelled (CA)13 and (GA)13 oligonucleotide repeats (Gautschi *et al*. 2000). Of 480 recombinant colonies screened, 86 gave a positive signal after hybridization. Plasmids from 69 positive clones were sequenced, of which primers were designed for 25 microsatellite inserts using PRIMER 3 (Rozen & Skaletsky 2000). After testing the primers' effectiveness on 12 individuals from different populations of the Mediterranean and redesigning primers for the ones that did not amplify well, nine loci were found to be polymorphic and amplify reliably.

Fluorescent-labelled primers were ordered for each locus; PET, NED and VIC (Applied Biosystems) and 5′–FAM (Microsynth). Three multiplex systems were developed to allow simultaneous amplification in one polymerase chain reaction (PCR): Aaran6, Aaran2_34 and Aaran2_38; Aaran9, Aaran17, Aaran25b and Aaran2_05; and Aaran2_25 and Aaran2_35b (Table 1). In a final volume of 6 μL, PCRs contained 3 μL Multiplex Master Mix by QIAGEN, 1.4 μL.

RNase-free water and 1 μL template DNA were diluted 1:1. To this mixture we added 0.6 μL of primer mix at the concentrations according to Table 1. PCR amplifications were conducted on a Whatman Biometra T1 Thermocycler as follows: 95 °C for 15 min; 27 cycles of 94 °C for 30 s, 54 °C for 3 min and 72 °C for 30 s; and a final extension of 72 °C for 30 min. Electrophoresis of the amplified products was performed using an ABI 3730 sequencer and scored with GENEMAPPER version 3.7 by Applied Biosystems.

We assessed genetic diversity for the nine loci in 25 individuals from Faro, Portugal, and 20 individuals from La Herradura, Spain. Non-amplifying samples were independently re-amplified for a second time.

Cross-species amplifications were conducted on 13 individuals of *Astropecten irregularis* from Muravera, Sardinia, two individuals of *Tethyaster subinermis* from Faro, Portugal, and three individuals of *Archaster* sp. from Fiji.

Within the two populations of *A. aranciacus,* the number of observed alleles ranged from four to 20. Using ARLEQUIN version 3.11 (Excoffier *et al*. 2005), expected and observed heterozygosities ranged from 0.593 to 0.936 and from 0.222 to 0.900 (Table 2). After correcting for multiple testing applying a Bonferroni correction (k = 9, adjusted α = 0.006), loci Aaran17, Aaran25b and Aaran2_34 deviated significantly from Hardy–Weinberg expectations in both populations, whereas loci Aaran2_35b and Aaran2_38 only deviated in one of the two examined populations (Table 2). Using MICRO-CHECKER version 2.2.3. (van Oosterhout *et al*. 2004), we estimated null allele frequencies (r; Table 2) and detected an excess of homozygotes over most size classes in Aaran17, Aaran25b, Aaran2_34, Aaran2_35b and Aaran2_38, favouring the hypothesis that null alleles are present at these loci. Tests of linkage disequilibrium for each pair of loci showed significant linkage in both populations for one pair (Aaran25b and Aaran2_34), and some pairs were significantly linked in one or the other population after Bonferroni correction. However, as significant linkage only occurred for pairs of loci where null alleles are likely to be present, we exclude the possibility of close physical linkage between any of the examined loci.

The cross-species amplification for *T. subinermis* was only successful at locus Aaran2_05 and amplified one allele with a length of 188 bp. Amplifications for *A. irregularis* resulted in five polymorphic loci with allele numbers ranging from three to nine. Six loci amplified for *Archaster* sp. of which four were polymorphic with two alleles each (Table 2).

The relatively moderate number of useful loci for *A. aranciacus* resulting from an initially promising set of positive clones indicates that isolation of suitable microsatellite loci for this group of echinoderms is not easy. Nevertheless, we are confident that the loci we characterize in this note are suitable to investigate the population structure of *A. aranciacus.* Furthermore, some of these loci might also serve for studies on other asteroids.

Acknowledgements

We would like to thank Teresa Cerveira Borges and the BIOPESCAS team (CCMAR, University of Algarve) for providing samples collected during the BIOFISH project. We thank Thomas Bucher, Andy Pemberton, Rob Toonen and Peter Wandeler for their technical support. Part of this project was funded by the Swiss National Science Foundation.

References

Burla H, Ribi G, Ferlin V, Pabst B (1972) Notes on Ecology of *Astropecten-Aranciacus*. Mar Biol **14**: 235-241.

Excoffier L, Laval G, Schneider S (2005) Arlequin ver. 3.0: An integrated software package for population genetics data analysis. Bioinformatics Online **1**: 47-50.

Gautschi B, Tenzer I, Muller JP, Schmid B (2000a) Isolation and characterization of microsatellite loci in the bearded vulture (*Gypaetus barbatus*) and cross-amplification in three Old World vulture species. Mol Ecol **9**: 2193-2195.

Gautschi B, Widmer A, Koella J (2000b) Isolation and characterization of microsatellite loci in the dice snake (Natrix tessellata). Mol Ecol **9**: 2191-2193.

Hörstadius S (1938) Über die Entwicklung von *Astropecten aranciacus* L. Pubblicazioni della Stazione Zoologica Napoli **17**: 221-312.

Rozen S, Skaletsky HJ (2000) Primer3. Code available at http://primer3.sourceforge.net/

Tenzer I, degli Ivanissevich S, Morgante M, Gessler C (1999) Identification of microsatellite markers and their application to population genetics of *Venturia inaequalis*. Phytopathology **89**: 748-753.

Van Oosterhout C, Hutchinson WF, Wills DPM, Shipley P (2004) MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. Mol Ecol Notes **4**: 535-538. doi: 10.1111/j.1471-8286.2004.00684.x.

Yasuda N, Nagai S, Hamaguchi M, Lian CL, Nadaoka K (2006) Development of microsatellite markers for the crown-of-thorns starfish *Acanthaster planci*. Mol Ecol Notes **6**: 141-143. doi: 10.1111/j.1471-8286.2005.01168.x.

Table 1 *Astropecten aranciacus* microsatellite loci characteristics and final primer concentration in Multiplex (*conc*). Accession numbers in GenBank are EU046580 - EU046588.

Locus	**Repeat motif**	**Primer sequence (5'-3')**	**conc (µM)**	**Multiplex panel**
Aaran06	$(CA)_{21}$	AGCCTGAACCCTTAGTGAGC **NED**-TTCCCTCAAACTGAAGTACCC	0.5	A
Aaran09	$(GT)_2A(CA)_{11}$	**NED**-TTGAGTGACAGCCGAATCAC CAGTATGCTATAGTCCGTACAACG	0.5	B
Aaran17	$(GT)_4AT(GT)_{32}$	**PET**-GCAAAACAAACTCTGTTGAGG GCTATCTCATTCGGACAGTGC	0.5	B
Aaran25b	$(CA)_4GCACG(CA)_{27}$	**FAM**-CGGTATGCAACGTACTACTG CAGTACATATTTGCTTTGAACC	2	B
Aaran2_05	$(TG)_{12}$	TCAGAATTTATATTAGCTGTGGATGC **FAM**-CCACGGTCAGGAACTACTGC	2	B
Aaran2_25	$(TG)_{32}$	GCACCTTCCTGTCGATCC **PET**-CGCGTGGATATAATTTACGG	2	C
Aaran2_34	$(CA)_4CGCACG(CA)_{26}$	**VIC**-CCGGTATGCAACGTACTACTG CTACAGTACATATTTGCTTTGAACC	2	A
Aaran2_35b	$(TG)_{10}AGATACAG(TG)_4$	**VIC**-CCAATGGTTTAAAGCTGTCC CACTAGGCACGGTAGATATGC	2	C
Aaran2_38	$(TG)_{35}$	GAGGTACCCCATAAGTGTCC **FAM**-CCAAGACGTTACTTCCAAGC	0.5	A

Table 2 Microsatellite allelic diversity within populations (*A*) and total allelic diversity among poplations (A_{tot}), observed (H_O) and expected (H_E) heterozygosity and estimate of null allele-frequency (*NA*) in *Astropecten aranciacus* samples of populations from Faro (*N*=25) and La Herradura (*N*=20). Additionally, cross-species amplification data for *A. irregularis* (*N*=13) and *Archaster* sp. (*N*=3) is presented for each locus.

Locus	**Population or species**	**No. of non-amplifying samples**	**Size range (bp)**	***A***	A_{tot}	H_O	H_E	***NA***
Aaran06	Faro	0	84-107	11	12	0.760	0.782	0.003
	La Herradura	0	88-109	10		0.850	0.844	-0.015
	A. irregularis	-	-	-				
	Archaster sp.	-	-	-				
Aaran09	Faro	0	117-128	5	5	0.720	0.736	-0.012
	La Herradura	0	117-123	4		0.750	0.717	-0.048
	A. irregularis	2	113-125	3				
	Archaster sp.	0	111	1				
Aaran17	Faro	5	114-168	12	16	0.300*	0.836	0.323
	La Herradura	2	114-178	10		0.222*	0.833	0.345
	A. irregularis	2	113-139	4				
	Archaster sp.	0	112	1				
Aaran25b	Faro	0	96-120	11	14	0.440*	0.860	0.267
	La Herradura	2	94-118	10		0.294*	0.886	0.319
	A. irregularis	5	108-125	3				
	Archaster sp.	0	83-115	2				
Aaran2_05	Faro	0	193-207	9	12	0.640	0.745	0.040
	La Herradura	0	195-211	10		0.900	0.833	-0.054
	A. irregularis	2	200-215	9				
	Archaster sp.	0	172-190	2				
Aaran2_25	Faro	2	74-143	20	25	0.826	0.934	0.051
	La Herradura	0	70-147	16		0.800	0.936	0.064
	A. irregularis	-	-	-				
	Archaster sp.	-	-	-				
Aaran2_34	Faro	0	102-125	11	14	0.480*	0.870	0.249
	La Herradura	2	96-123	11		0.333*	0.883	0.300
	A. irregularis	2	111-118	6				
	Archaster sp.	1	115-142	2				
Aaran2_35b	Faro	4	113-151	9	10	0.286*	0.774	0.312
	La Herradura	4	113-141	7		0.438	0.593	0.107
	A. irregularis	-	-	-				
	Archaster sp.	0	145-173	2				
Aaran2_38	Faro	5	125-200	15	22	0.600	0.874	0.186
	La Herradura	5	134-215	17		0.600*	0.934	0.171
	A. irregularis	-	-	-				
	Archaster sp.	-	-	-				

* indicates significant deviation from Hardy-Weinberg equilibrium (K=9, α=0.006).

- no amplification.

CHAPTER IV

Genetic structure of the high dispersal Atlanto-Mediterreanean sea star *Astropecten aranciacus* revealed by mitochondrial DNA sequences and microsatellite loci

Deborah E. Zulliger, Samuel Tanner, Markus Ruch and Georg Ribi

Abstract

To investigate the impact of potential marine barriers on gene flow in high dispersal marine invertebrates, we assessed the population genetic structure of the sea star *Astropecten aranciacus*. Samples were obtained from nine locations within the Atlantic and the Mediterranean Sea including populations east of the Siculo-Tunisian Strait. We obtained both DNA sequence data of the mitochondrial control region and genotype data at four microsatellite loci. Both markers were highly polymorphic and showed a great level of genetic diversity. Genetic differentiation between populations (F_{ST}) was in general low, particularly for nuclear data, as is often the case in high dispersal marine invertebrates. Nevertheless, both marker sets indicated a significant genetic differentiation of the population from the island of Madeira to most other populations. Our results also demonstrate a clear pattern of isolation-by-distance supported by both mitochondrial and nuclear markers. Therefore, we conclude that larval dispersal of *A. aranciacus* is somewhat limited even within the basins of the Atlantic, the west Mediterranean and the east Mediterranean. Microsatellite loci further revealed genetic differentiation between the three basins; however, it is not clear whether this is truly caused by marine barriers. Genetic differentiation between basins might also be a result of isolation-by-distance allowing for any grouping to be significant as long as geographical neighbors are clustered together. Although levels of genetic differentiation were less pronounced in microsatellite data, both datasets were coherent and revealed similar patterns of genetic structure in *A. aranciacus*.

Introduction

Marine species with extended planktonic larval stages have a high capacity for dispersal and as such are expected to display less genetic structure than species without a long stage in the plankton (Palumbi and Wilson 1990). Nevertheless, gene flow in marine species can be constrained by dispersal barriers, such as narrow water passages between land masses, sharp salinity gradients or different types of currents e.g., circular currents (eddies) or downward currents. As marine barriers are not always easily identified, they might lead to population structure even in high dispersal species (Quesada et al. 1995; Palumbi et al. 1997).

The role of the Atlantic-Mediterranean division as a potential barrier to gene flow has increasingly been investigated for various planktotrophic invertebrate species (e.g., Borsa et al. 1997; Launey et al. 2002; Diaz-Almela et al. 2004; Duran et al. 2004a; Stamatis et al. 2004, 2006; Saavedra and Pena 2005; Calderon et al. 2008). While the Strait of Gibraltar geographically divides the two basins, the Almería-Oran front is thought to be a genetic separation area. This large-scale density front is formed by the convergence of two distinct water masses in the east Alboran Sea, which is located in the westernmost region of the Mediterranean Sea (Tintore et al. 1988). Other barriers to gene flow within the Mediterranean may also exist in the form of an east–west divide at the Siculo-Tunisian Strait and/or hydrogeographic isolation of the Aegean, Ionian and Adriatic Seas (Perez-Losada et al. 2007).

Many studies have used indirect genetic tools such as mitochondrial DNA (mtDNA), nuclear DNA or a combination of the two to analyze genetic structure in high dispersal Atlanto-Mediterranean invertebrates (e.g., Féral et al. 1995; Zane et al. 2000; Launey et al. 2002; Diaz-Almela et al. 2004; Duran et al. 2004a; Roman and Palumbi 2004; Stamatis et al. 2004; Triantafyllidis et al. 2005; Peijnenburg et al. 2006; Calderon et al. 2008). Differing conclusions regarding the influence of the Atlantic–Mediterranean division on population structuring were drawn in these studies depending not only on the species investigated but also on the genetic markers used and the sampling pattern. For instance, only moderate genetic differentiation was revealed between Atlantic and Mediterranean populations of the sea urchin *Paracentrotus lividus* based on mtDNA sequences (Duran et al. 2004a), whereas a sharp break was detected between the two basins when combining mitochondrial and nuclear markers and applying a more extensive sampling (Calderon et al. 2008). Based on allozymes and 28S rRNA sequence data, a clear separation was also detected in the sea urchin *Echinocardium cordatum* (Féral et al. 1995). Patarnello et al. (2007) discovered that even between closely related taxa with comparable biologies the Atlanto-Mediterranean transition does not always induce a congruent population genetic structure, which could be due to the differences in demographic history between these species. Restriction fragment length polymorphism (RFLP) of mtDNA in the Norway lobster *Nephrops norvegicus*, for example, showed no genetic differentiation between the Atlantic and the Mediterranean (Stamatis et al. 2004). However, pronounced differentiation between Atlantic and Mediterranean populations was detected in the closely related European lobster *Homarus gammarus* (Triantafyllidis et al. 2005) using also RFLP of mtDNA. Two other crustaceans, the high dispersal green crab

Carcinus maenas (Roman and Palumbi 2004) and the pelagic Northern krill *Megenyctiphanes norvegica* (Zane et al. 2000), again showed genetic differentiation between the basins based on mtDNA sequences.

Gene flow in marine species with long planktonic larval stages may be more restricted than generally assumed. In such cases, random drift occurs locally, and genetic structure can develop in the form of isolation-by-distance. This genetic pattern has been revealed in some high dispersal marine invertebrates, such as in the European flat oyster *Ostrea edulis* using microsatellite loci and mtDNA sequence data (Launey et al. 2002; Diaz-Almela et al. 2004) and in the pelagic crustacean *Megenyctiphanes norvegica* (Zane et al. 2000) using mtDNA data only.

Here we present data from another high dispersal Atlanto-Mediterranean echinoderm: the sand star *A. aranciacus*. *A. aranciacus* is a broadcast spawning sea star, which undergoes a long planktotrophic larval stage. Due to its large body size of up to 60 cm in diameter and the potential for high population densities, it is believed to be an important benthic predator (Burla et al. 1976). In the Mediterranean, this sea star was once abundant (Burla et al. 1972). However, over the past 20 years a decline in populations of *A. aranciacus* has been observed in several areas within the Mediterranean (G. Ribi unpublished; H. Lessios, H. Massé, H. Moosleitner, L. Santella, personal communications). The present distribution of *A. aranciacus* includes the Mediterranean Sea and the east Atlantic coast from northern Portugal to Angola, including the Canary Islands, Cape Verde and Madeira (e.g., Koehler 1921; Tortonese 1980). This sea star usually lives in depths of 1-100 m (Zavodnik 1960), but has been found at depths of up to 183 m (Hörstadius 1938). Migration of adult sea stars is therefore bound to the continental shelf and does not tend to occur in any particular direction (Pabst 1986); hence, adult movements are likely to be only a minor factor for dispersal. In southern Portugal, *A. aranciacus* is still highly abundant (C. Almeida, personal communication). With a planktonic larval stage of up to 60 days (Hörstadius 1938), *A. aranciacus* larvae can likely disperse up to 400 km following the calculations of Shanks et al. (2003). This high potential for dispersal might allow populations in the east Atlantic, such as for instance from southern Portugal, to replenish the Mediterranean populations along the prevailing water exchange direction, if there are no barriers to larval dispersal between these two basins.

While sea stars have been subject to several population genetic studies using mitochondrial and/or nuclear markers (Hunt 1993; Williams 2000; Williams and Benzie 1997, 1998; Williams et al. 2002; Matsuoka and Asano 2003; Waters et al. 2004; Waters and Roy 2004; Colgan et al. 2005; Harper and Hart 2005; Harley et al. 2006; Harper et al. 2007; Gerard et al. 2008), only one study was conducted in the Atlanto-Mediterranean region (Baus et al. 2005). This study found high genetic structure between Atlantic and Mediterranean populations of *Asterina gibbosa*, a sea star which is expected to have a low dispersal capacity, as it lacks a planktotrophic larval stage. The present study investigates potential marine barriers to gene flow in a high dispersal marine invertebrate by analyzing the population genetic structure of *A. aranciacus* employing both mitochondrial and nuclear markers. The comparison of these two marker types allows a more comprehensive investigation of genetic diversity, as markers of these two physically unlinked genomes do not always show congruent patterns (e.g., Hansen et al. 1999; Lemaire et al. 2005; Costantini et

al. 2007). In this study, we sequenced the complete control region of the mitochondrial DNA, a region that has the highest rate of evolutionary change of any mtDNA region (Aquadro and Greenberg 1983; Parsons et al. 1997). As nuclear markers, we used four polymorphic microsatellite loci, which are believed to be neutral and have been shown to be more variable compared to e.g., allozyme data (Shaw et al. 1999; Estoup et al. 1998; Perez-Losada et al. 2002).

Employing these methods we (1) investigate a possible genetic separation between *A. aranciacus* populations of the Atlantic and the Mediterranean basin and between populations in the eastern and the western Mediterranean; (2) test existing populations for isolation-by-distance versus panmixia within and/or among the basins; and (3) determine the degree of correlation between genetic differentiation patterns estimated from mtDNA sequence data and patterns resulting from microsatellite loci.

Materials and methods

Sampling and molecular methods

We sampled a total of 254 individuals from 9 locations within the Mediterranean Sea and the east Atlantic as shown in Fig. 1. Specimens were obtained by scuba diving and from commercial trawl and gill net operations within the years 2002 to 2006. Samples were preserved in 96% ethanol or in 80% ethanol buffered with DMSO until processed. We extracted DNA of approximately 30 mg of arm tip tissue or tube feet using a DNeasy Tissue Kit® (QIAGEN) following the manufacturer's instructions for extraction of animal tissue for a final volume of 400 µL. Extracted DNA was stored at -20°C.

We amplified fragments of the mitochondrial DNA (mtDNA) for 15-20 specimens per location according to Table 1 using the forward primer E12Sa and the reverse primer E16Sb as described in Smith et al. (1993). These primers amplify a fragment of approximately 1,200 base pairs (bp) in *A. aranciacus* and contain part of the 16S ribosomal RNA region, the entire non-coding control region (CR), the tRNA-Thr/Glu regions and part of the 12S ribosomal RNA region. DNA amplifications were performed in 30 µL -volume reactions with 1.67 U *Taq* DNA *Polymerase*, 3 µL 10x PCR reaction buffer, 0.4 mM dNTPs, 0.2 µM of each primer, 1 mM MgCl2 and 6 µL of DNA. The PCR protocol consisted of an initial denaturation step at 95°C for 3 min, 40 amplification cycles (95°C for 30 s, 48°C for 30 s and 72°C for 1 min) and a final elongation step at 72°C for 10 min performed in a Whatman Biometra T1 Thermocycler. We purified the PCR products with the QIAquick® PCR Purification Kit (Qiagen) or NucleoSpin® Extract II (Macherey-Nagel AG, Oensingen, Switzerland), following the supplier's instructions. Forward and reverse sequencing using E12Sa and E16Sb were carried out separately using BigDye® Terminator (PE-Applied Biosystems) chemistry. The cycle-sequencing protocol consisted of an initial step at 96°C for 3 min and 24 sequencing cycles (96°C for 15 s, 50°C for 10 s and 60°C for 3 min). Cycle sequencing products were purified with a DyeEx™ 2.0 Spin Kit (Qiagen) or a NucleoSeq® (Macherey-Nagel) Purification Kit and sequenced on an ABI 3730 DNA

Analyzer. Sequences were edited and aligned using the software SEQUENCHER™ 3.0 (Gene Codes Corporation) and adjusted by eye.

We analyzed nuclear DNA variation for all samples at four polymorphic microsatellite loci (Aaran06, Aaran09, Aaran2/05 and Aaran2/25) as previously characterized for *A. aranciacus* by Zulliger et al. (2008). DNA amplifications, fragment analyses and scoring were performed as described by Zulliger et al. (2008).

Genetic diversity within populations

MtDNA haplotypic diversity h (Nei 1987) and nucleotide diversity π (Tajima 1983) were calculated using ARLEQUIN version 3.11 (Excoffier et al. 2005). Moreover, we constructed a haplotype network using the statistical parsimony procedure of Templeton et al. (1992) implemented in TCS version 1.21 (Clement et al. 2000) with gaps coded as 5th character state and a 95% connection limit. To test for possible nucleotide saturation, we obtained saturation plots using the program DAMBE (Data Analysis in Molecular Biology and Evolution; Xia and Xie 2001).

Microsatellite allelic diversity (N_a), allele frequencies and allelic richness corrected for differences in samples size (R_S; ElMousadik and Petit 1996) were determined per locus and per sampling location using the software FSTAT version 2.9.3.2 (Goudet 1995). Further, we calculated observed (H_O) and expected heterozygosity (H_E) in ARLEQUIN and tested for significant deviations from Hardy-Weinberg equilibrium (HWE) as described in Guo and Thompson (1992) using 1,000,000 steps in the Markov chain and 50,000 dememorization steps. Tests for linkage disequilibrium were also performed in ARLEQUIN using a likelihood-ratio test (Slatkin and Excoffier 1996) and 16,000 random permutated samples. To detect microsatellite scoring errors, large allele dropout, occurrence of null alleles and estimates of null allele frequency, we used the software MICRO-CHECKER version 2.2.1 (Van Oosterhout et al. 2004).

As the haplotype diversity (h) and the allelic richness (R_S) were noticeably lower in Madeira (MAD) than in the other locations, we performed a one-sample t test using the statistics software SPSS 14.0 to test for statistical significance of this difference.

Bayesian clustering

STRUCTURE version 2.0 (Pritchard et al. 2000) was used to infer population genetic structure testing the consistency with microsatellite genetic information. This Bayesian clustering method takes a sample of genotypes and uses the assumption of HWE and linkage equilibrium within subpopulations to find the number of populations (K) that fits the data best and the individual assignments that minimize Hardy-Weinberg and linkage disequilibrium in those populations. 10 replicates of this analysis were performed with K ranging from 1 to 12 for 1,000,000 generations (burn-in 100,000) and assuming an admixture model.

Genetic differentiation among populations

Population pairwise F_{ST} estimates (Weir and Cockerham 1984; Michalakis and Excoffier 1996) were calculated using ARLEQUIN for mtDNA and microsatellite data applying 16,000 permutations and Kimura 2-parameter corrected distances for mtDNA. Estimates of F_{ST} based on microsatellite data were also carried out without locus Aaran2/25, as null alleles are likely to be present at this locus (see below). To test whether F_{ST} using all four loci and F_{ST} without locus Aaran2/25 differed significantly, a paired sample *t* test was performed with the statistics software SPSS. This test was not significant (P = 0.648, correlation = 0.948), and thus, all further calculations based on F_{ST} were carried out using all four loci. For microsatellite data we also calculated the standardized genetic differentiation measure as proposed by Hedrick (2005) which accounts for the level of genetic variation. To calculate this measure, all alleles were recoded as being population specific using the program RECODEDATA (Meirmans 2006).

We examined the partitioning of the total variance between various groups of samples in ARLEQUIN by performing hierarchical analysis of molecular variance (AMOVA; Excoffier et al. 1992). Based on allelic frequencies for microsatellite loci and applying Kimura 2-parameter corrected distances for mtDNA, multilevel AMOVAs were performed to examine the proportion of genetic variance among the Atlantic and Mediterranean basins and within the Mediterranean among the west and the east basins separated by the Siculo-Tunisian strait (see Fig. 1). As the sampling location La Herradura (LAH) is located in the Alboran Sea, west of the Almería-Oran front, AMOVAs were carried out both with LAH belonging to the Atlantic and with LAH belonging to the west Mediterranean (16,000 permutations).

To estimate the effects of isolation-by-distance, we calculated the correlation between pairwise genetic differentiation (F_{ST}) and geographic distance using the software IBD version 1.52 (Bohonak 2002). IBD uses a Mantel-test to find relationships between genetic and geographic distance matrices. As proposed by Rousset (1997) for a two-dimensional dispersal, we compared F_{ST} /(1 - F_{ST}) with the logarithm of the geographic distance. Isolation-by-distance was tested for distances measured as the direct sea path between sampling locations (10,000 randomizations). As pairwise genetic differentiation (F_{ST}) of Madeira (MAD) to most other locations was significant, we performed a second analysis without this location to test whether the correlation was an artefact caused by the high genetic differentiation to MAD.

Correlation between mtDNA and microsatellite F_{ST}

We determined the correlation between the two matrices of population pairwise genetic differentiation (F_{ST}) resulting from mtDNA and microsatellite loci using a Mantel-test as applied in IBD (10,000 randomizations).

When multiple tests were performed, we corrected the level of significance according to the number of tests in a given set applying the control of false discovery rate (FDR) method (Benjamini and Hochberg 1995) as suggested by Narum (2006).

Results

Genetic diversity within populations

We successfully amplified a 1,017 bp fragment of mtDNA containing the control region in 151 individuals across eight Atlantic and Mediterranean sampling locations (see Fig. 1, Table 1). All sequences were deposited in GenBank under accession numbers EU450469–EU450582. The sample from Banyuls (BAN) was not included in the mtDNA analyses, because it was not possible to amplify the desired fragment of enough individuals from this location.

A total of 114 unique haplotypes were identified of which the two most common haplotypes (no. 1 and 72) were found in all three regions (see Appendix for absolute haplotype frequencies). We observed 96 polymorphic sites (9.4% variable sites), of which 12 contained gaps, and the number of nucleotide substitutions between any pair of sequences ranged from 1 to 25. Mean haplotype diversity (h) equalled 0.99 with the highest h observed in the sample from Muravera (MUR; h = 1.00) and the lowest h in Madeira (MAD; h = 0.84). A one-sample t test revealed a significantly lower h in MAD compared to other populations (mean difference = 13.385; $P < 0.001$). Mean nucleotide diversity (π) was 0.0064, and no evidence of nucleotide saturation was present in the saturation plot produced by DAMBE (results not shown).

One minimum spanning haplotype network with numerous ambiguous connections was obtained using TCS (Fig. 2). A central haplotype could be identified in all three regions, and the most frequent haplotypes were generally closely linked to each other.

For the majority of the individuals, all four microsatellite loci amplified successfully and were scored unambiguously (Table 1). At loci Aaran06 and Aaran2/05 one individual each did not amplify, and at locus Aaran2/25 amplifications for 13 individuals from five different sampling locations were not successful. All loci were polymorphic at each sampling location, and allelic diversity (N_a) within populations ranged from three alleles in Aaran09 to 28 alleles in Aaran2/25 (Table 1). There was no evidence of scoring errors due to stutter or large allele dropout. Observed (H_O) and expected heterozygosity (H_E) ranged from 0.429 to 0.933 and from 0.547 to 0.968, respectively. Except for locus Aaran2/25, none of the loci showed evidence of the presence of null alleles. Deviations from Hardy-Weinberg equilibrium (HWE) were detected at locus Aaran2/25 after correcting for multiple testing (initial $\alpha = 0.05$, k = 4) for the samples Madeira (MAD), Faro (FAR), La Herradura (LAH), Banyuls (BAN), Muravera (MUR), Kavala (KAV) and for all samples combined. The presence of one or more null alleles at locus Aaran2/25 was indicated by an excess of homozygotes over most size classes, detected by the program MICRO-CHECKER, and by the failure to amplify this locus in several specimens. In contrast, a significant excess of heterozygotes was detected at locus Aaran09 within the population MUR. Evidence of linkage disequilibrium between pairs of loci among populations was only significant at Gaeta (GAE) for loci Aaran06 and Aaran2/25 (P = 0.005), while over all populations no loci were significantly linked. Allelic richness adjusted for differences in sample size ranged from 3.000 in MAD at Aran09 to 23.131 in KAV at Aaran2/25 (Table 1). Overall allelic richness as an

average of the four loci was the highest in Heraklion (HER) with 12.054 and the lowest in MAD with 8.975 (Table 1). A one-sample *t* test showed a significantly lower allelic richness in MAD compared to the other locations (mean difference = 2.180; $P < 0.001$). Over all microsatellite loci, 18.7% of the alleles were population specific, whereas the lowest percentage of population specific alleles equalled 15.2% at locus Aaran2/25. The number of population specific alleles ranged from two at Aaran09 to seven at Aaran2/25 and amounted to a total of 17 alleles over all loci.

Bayesian clustering

Structure analysis of microsatellite data failed to distinguish among the nine locations, as all ten runs exhibited the best -log PR($X|K$) estimates for $K = 1$ (lnL = -4273.5; Fig. 3). We repeated this analysis without locus Aaran2/25, since this locus demonstrated deviation from HWE, however, the best -log PR($X|K$) estimates remained for $K = 1$ (5 runs; lnL = -2781.7; results not shown).

Genetic differentiation among populations

MtDNA pairwise F_{ST} ranged from 0 to 0.19 and provided evidence of a genetic subdivision between MAD and all other locations (Table 2). Microsatellite loci revealed a similar genetic differentiation pattern, but in addition to MAD, F_{ST} were also significant between FAR and several Mediterranean populations. F_{ST} in microsatellites ranged from zero to 0.0312 and were on average 5.9-fold lower than in mtDNA (Table 2). Nevertheless, a Mantel-test showed significant correlation between mitochondrial and microsatellite F_{ST} ($r = 0.84$; $P = 0.0127$). Standardized genetic differentiation in microsatellite data ranged from 0 to 0.15 and was on average 1.5-fold lower than in mtDNA.

Analyses of molecular variance (AMOVA) revealed high levels of variation within populations for both mtDNA and microsatellites ranging from 94.15 to 97.21% and from 98.75 to 99.69%, respectively (Table 3). In mtDNA, variation among populations was the highest when comparing the west Mediterranean (without LAH) versus the east Mediterranean populations (3.94%). However, none of the groupings showed significant among group variation in mtDNA after correcting for multiple testing. In microsatellites, the highest percentage of among group variation was achieved when clustering the Atlantic versus the Mediterranean populations (0.8%; Table 3). Control of false discovery rate (FDR) left all groupings significant, except for one (WMed without LAH vs. EMed).

Isolation-by-distance analysis by the nearest sea path resulted in a significant correlation of genetic differentiation and geographic distance in both mtDNA ($P = 0.0043$; Fig. 4a) and microsatellite loci ($P = 0.0023$; Fig. 4b). Even when omitting the sample MAD from the analysis, this correlation remained significant in both markers ($P_{mtDNA} = 0.0261$; $P_{msat} = 0.0038$). As shown in Fig. 4b, pairs with MAD (empty dots) clearly have higher F_{ST} than could be expected by the geographic distance, as well as one outlier pair with FAR and BAN (filled dot).

Discussion

In this study we used both mitochondrial and microsatellite markers to investigate the genetic structure in a high dispersal echinoderm in the Atlanto-Mediterranean region. Our results showed that the mitochondrial control region is highly variable and the four microsatellite loci Aaran06, Aaran09, Aaran2/05 and Aaran2/25 are also highly polymorphic in *A. aranciacus*. The comparison of these two markers allows us to determine the population structure in this sea star in a comprehensive manner, as our mitochondrial and nuclear data in general exhibit a congruent pattern of genetic differentiation.

Within population variability

While nucleotide diversity was low in mtDNA sequence data, haplotype diversity was high, indicating a high degree of polymorphism in the mitochondrial control region. Other studies on echinoderms have obtained similar results for sequences of the mitochondrial cytochrome c oxidase subregion I (COI) (e.g., McCartney et al. 2000; Uthicke and Benzie 2003; Duran et al. 2004a). In marine invertebrates with large population sizes numerous haplotypes can be retained during periods of population growth or expansion (Watterson 1984). A rapid population expansion could therefore lead to high haplotype and low nucleotide diversity as new mutations are retained (Avise et al. 1984; Watterson 1984). In the present case, the results of the parsimony network analysis for *A. aranciacus* seem to support this hypothesis. The network showed a star-shaped pattern with a central haplotype and various alternative connections. The star-shape possibly indicates a common ancestral haplotype (Templeton et al. 1995), supported by the occurrence of a central haplotype in all three basins (Atlantic, western and eastern Mediterranean).

Microsatellite loci were also highly polymorphic and showed a high mean allelic richness per population (R_s = 21.9), but only few alleles were population specific. The lowest percentage of population specific alleles was found at locus Aaran2/25 (15.2%). Together with the high allelic richness at this locus, this could be an indication of size homoplasy. According to Estoup et al. (2002), this would not necessarily be a significant problem for many types of population genetic analyses. However, highly polymorphic loci due to high mutation rates can lead to lower estimates of F_{ST} (Slatkin 1995), and high levels of heterozygosity may reduce the relationship between statistical and biological significance (Hedrick 1999). As F_{ST} estimates with and without locus Aaran2/25 showed no significant difference, we can assume that neither homoplasy nor the presence of null alleles skew the results of this study. Moreover, multiallelic F_{ST} is based on a weighted mean of the contribution to the overall variance of each allele considered separately (Weir and Cockerham 1984). It is therefore not likely to be biased by an invisible allele, which is expected to be randomly associated with size scored alleles (Launey et al. 2002).

For some groups of invertebrates the proportion of microsatellite loci without null alleles is often low (McGoldrick et al. 2000; Launey et al. 2002; Peijnenburg et al. 2006; Costantini et al. 2007). Possibly, this could be due to a less effective DNA repair mechanism in the nuclear DNA of some invertebrates, especially

in echinoderms. Primers for amplification of microsatellite loci have been developed for several other echinoderms, such as *Acanthaster plancii* (Yasuda et al. 2006), *Amphipholis squamata* and *Echinocardium cordatum* (Chenuil et al. 2003), *Strongylocentrotus spp.* (Addison and Hart 2002), *Parastichopus californicus* (Nelson et al. 2002), *Apostichopus japonicus* (Zhan et al. 2007), *Tripneustes gratilla* (Carlon and Lippe 2007) and *Evechinus chloroticus* (Perrin and Roy 2000).

Haplotype and allelic diversity are significantly lower in the sample from Madeira (MAD) than in all other samples, which could reflect either a founder event during the colonization of Atlantic islands or a recent bottleneck. The population from MAD was most likely not closely linked to the coastal populations in the past, as the island of Madeira emerged volcanically at the most 5 Mio years ago (Geldmacher and Hoernle 2000). Lower allelic richness has also been found in populations of *Crambe crambe* in the Canaries and Madeira archipelagos compared to Mediterranean populations (Duran et al. 2004b). Along with a reduction of alleles, an excess of heterozygosity would be a sign of a recent bottleneck, as allele number usually decreases faster than heterozygosity (Cornuet and Luikart 1996). We tested for a recent reduction in population size applying the software BOTTLENECK version 1.2.02 (Cornuet and Luikart 1996), but could not detect a bottleneck for MAD. Nevertheless, this possibility can not be completely ruled out, and additional microsatellite loci would be necessary to gain more clarity. Furthermore, genetic patterns are similar both in founder events and recent bottlenecks, and thus both possibilities still remain. Using mtDNA we performed a mismatch distribution analysis and Tajima's D test of selective neutrality (Tajima 1989) in ARLEQUIN to explore the demographic past of *A. aranciacus* in Madeira. This test showed that the population in MAD has undergone a recent population expansion, as it did not differ significantly from the sudden expansion model by Rogers and Harpending (1992).

Spatial structure

While Bayesian clustering (STRUCTURE) of microsatellite data failed to detect any genetic structure in *A. aranciacus*, pairwise genetic differentiation (F_{ST}) was significant mainly between MAD and all the other samples. These results are not contradictory given the low estimates of F_{ST} in microsatellites and that STRUCTURE tries to assign individuals to a population without any initial information about the true sampling location. Mitochondrial and nuclear data in general showed congruent genetic differentiation patterns. The significant genetic differentiation between MAD and the other populations revealed by mitochondrial and microsatellite data can be explained by the remoteness of MAD to the other sites and possibly also by currents which are unfavourable to movement from the shelf to Madeira and vice versa. The Portugal current, which flows along the east Atlantic coast from northern Portugal southwards to northern Africa, is a possible marine barrier to gene flow from the Mediterranean and east Atlantic coast to the island of Madeira. Moreover, the Mediterranean out-flow tends to stratify in the Atlantic Ocean at a depth of 600–1,400 m due to its greater density (Mougenot and Vanney 1982) and might impede larval dispersal to Madeira. On the

other hand, larvae dispersing from Madeira are prone to be directed southward along with the Canary current, restricting gene flow to the east Atlantic coast and the Mediterranean.

Although highly correlated, F_{ST} (θ) were several times higher in mitochondrial than in nuclear markers. The average cytoplasmic/nuclear ratio (θ_C/ θ_N) equalled 5.9. Similar results have been found in several other studies comparing mtDNA and microsatellite data (e.g., Shaw et al. 1999; Krafsur 2002; Diaz-Almela et al. 2004; Lemaire et al. 2005; Peijnenburg et al. 2006). As the effective mitochondrial population size (N_e) is expected to be four times smaller than the nuclear N_e, assuming both 1:1 sex ratios and diploidy (Birky et al. 1983), there is a higher potential for genetic drift in mtDNA, which leads to a higher θ. Ratios significantly higher than fourfold can be caused by sex bias in migration, reproduction or population size or by a difference in the mutation rate of the two markers (Turan et al. 1998; Shaw et al. 1999). The average θ_C/ θ_N ratio in *A. aranciacus* is slightly higher than fourfold, for which the mentioned possibilities of sex bias or differences in mutation rates, combined or separately, could be partly responsible. A sex bias in migration though is an improbable factor, given the non-migratory behaviour of *A. aranciacus*. The standardized measure of genetic differentiation for microsatellite data (Table 2) is almost comparable to F_{ST} in mtDNA and shows that low values of F_{ST} in microsatellite loci are mostly due to the high polymorphism in these markers.

Mitochondrial data failed to indicate any significant among group variation for any of the groups formed for the analysis of molecular variance (AMOVA) after correcting for multiple testing. In microsatellite loci, however, significant among group variation was found for all groupings except for one, indicating some differentiation between the basins. Since the Atlantic versus Mediterranean differentiation might be skewed by the sample from Madeira (MAD), further sampling along the east Atlantic coast, e.g., North Africa, is needed in order to clarify whether there is a true genetic differentiation between these two basins, or if the pattern observed here is limited to offshore islands. As for the Mediterranean, our data suggest significant among group variation between populations of *A. aranciacus* in the west and the east Mediterranean basins. A similar differentiation between eastern (Adriatic) and western Mediterranean populations has been found in other high dispersal marine species, such as the chaetognath *Sagitta setosa* (Peijnenburg et al. 2006) and the bivalve *Cerastoderma glaucum* (Mariani et al. 2002). Although AMOVA indicate some genetic differentiation between basins in populations of *A. aranciacus*, this variation could be connected to the clear isolation-by-distance pattern found in *A. aranciacus*, as grouping populations by basin includes a strong geographical component. This hypothesis deserves further consideration, as in most cases significance of among group variation did not depend on whether LAH was assigned to the Mediterranean or to the Atlantic group.

Analyses of isolation-by-distance (IBD) were significant in both mtDNA and microsatellite data regardless of whether MAD was included in the analysis or not. We can therefore rule out a possible artefact caused by MAD, which showed a significant genetic differentiation (F_{ST}) to most other populations. However, pairwise F_{ST} comparisons between MAD and the other study sites are much higher than expected from the regression line resulting from IBD analyses (see Fig. 4). In fact, F_{ST} values between MAD and the

other populations are about five times higher in mtDNA and twice as high in microsatellites. Besides the large geographic distance, other isolation mechanisms must therefore be acting on the population of MAD, such as the unfavourable current regime discussed above. Why the pair FAR-BAN also exhibits such a high F_{ST} in microsatellites still needs to be further investigated by either obtaining comparable mtDNA data from BAN or by sampling other populations of the west coast of the Mediterranean. Isolation-by-distance patterns have been detected in other Atlanto-Mediterranean invertebrates (Zane et al. 2000; Launey et al. 2002; Diaz-Almela et al. 2004) but not yet in high dispersal echinoderms. In contrast to our results, genetic investigations on the high dispersal sea urchin *Paracentrotus lividus* (Duran et al. 2004a) suggested panmixia within the Mediterranean and within the east Atlantic basin. This discrepancy is most likely due to the sampling scheme, as the presented study included populations in the eastern Mediterranean while the sampling by Duran et al. (2004a) was limited to the west coast of the western Mediterranean. While mitochondrial and nuclear datasets reveal similar levels of significant genetic differentiation between samples, and both support an isolation-by-distance pattern, mtDNA showed higher population differentiation (F_{ST}) than microsatellite loci. However, when accounting for the high polymorphism in nuclear markers using the standardized genetic differentiation measure, this difference was less pronounced. On the other hand, microsatellite loci were more sensitive in detecting genetic differentiation between groupings by basins (AMOVA).

Conclusions

Our results indicate that while genetic differentiation may be mostly absent in *A. aranciacus*, the dispersal of marine invertebrates with extended planktonic larval stages could be more restricted than is often assumed. In contrast to previous studies on echinoderms suggesting panmixia in the Atlantic and Mediterranean basin, our data revealed a pattern of isolation-by-distance in *A. aranciacus* over the sampled area. Microsatellite data further detected some differentiation of Atlantic versus Mediterranean and western versus eastern Mediterranean populations of *A. aranciacus*. Nevertheless, our data did not allow to identify specific marine barriers, such as the Strait of Gibraltar, the Almería-Oran front or the Siculo-Tunisia Strait, as isolation-by-distance might be sufficient to explain the majority of the genetic differences found here. Further sampling, particularly along the Atlantic coast, is necessary to gain more clarity on this matter. The present study highlights how the comparison of mitochondrial and microsatellite markers can provide a more complete picture of genetic differentiation, allowing for more comprehensive data analyses and interpretation.

Acknowledgments

We thank Luigia Santella, Teresa Cerveira Borges and the BIOPESCAS team (CCMAR, University of Algarve), Paulo Morais, Martin von Arx, Heinz Maag and Till Danckwart for providing specimens from Gaeta, Naples, Faro, Banyuls, Muravera and Cres. Santi Diliberto and Susanna Tassis enabled further sampling in Greece. Thomas Bucher, Andy Pemberton, Heinz Maag and Marco Bernasconi all provided lab and technical support. We are grateful for the assistance with data analyses that we received from Peter Wandeler, Tony Wilson and Rob Toonen and for helpful suggestions from anonymous reviewers. This project was funded in part by the Swiss National Science Foundation.

References

Addison JA, Hart MW (2002) Characterization of microsatellite loci in sea urchins (*Strongylocentrotus* spp.). Mol Ecol Notes 493–494. doi:10.1046/j.1471-8286.2002.00295.x.

Aquadro CF, Greenberg BD (1983) Human mitochondrial-DNA variation and evolution-analysis of nucleotide-sequences from 7 individuals. Genetics **103**:287–312.

Avise JC, Neigel JE, Arnold J (1984) Demographic influences on mitochondrial-DNA lineage survivorship in animal populations. J Mol Evol **20**: 99–105. doi:10.1007/BF02257369.

Baus E, Darrock DJ, Bruford MW (2005) Gene flow patterns in Atlantic and Mediterranean populations of the Lusitanian sea star *Asterina gibbosa*. Mol Ecol **14**: 3373–3382. doi: 10.1111/j.1365294X.2005.02681.x.

Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate - a practical and powerful approach to multiple testing. J R Stat Soc B Methodol **57**: 289–300.

Birky CW, Maruyama T, Fuerst P (1983) An approach to population and evolutionary genetic theory for genes in mitochondria and chloroplasts, and some results. Genetics **103**: 513–527.

Bohonak AJ (2002) IBD (isolation by distance): a program for analyses of isolation by distance. J Hered **93**: 153–154. doi:10.1093/jhered/93.2.153.

Borsa P, Blanquer A, Berrebi P (1997) Zoogéographie intraspécifique de la mer Méditerranée. Analyse des données génétiques populationnelles sur seize espèces atlanto-méditerranéennes (Poissons et Invertèbres). Vie Milieu **47**: 95–305.

Burla H, Pabst B, Stahel W (1976) Environmental-conditions affecting occurrence of *Astropecten-Aranciacus* (Asteroidea, Echinodermata). Helgol Wiss Meeresunters **28**: 167–182. doi: 10.1007/BF01610351.

Burla H, Ribi G, Ferlin V, Pabst B (1972) Notes on ecology of *Astropecten-Aranciacus*. Mar Biol (Berl) **14**: 235.

Calderon I, Giribet G, Turon X (2008) Two markers and one history: phylogeography of the edible common sea urchin *Paracentrotus lividus* in the Lusitanian region. Mar Biol (Berl) **154**: 137–151. doi: 10.1007/s00227-008-0908-0.

Carlon DB, Lippe C (2007) Eleven new microsatellite markers for the tropical sea urchin *Tripneustes gratilla* and cross-amplification in *Tripneustes ventricosa*. Mol Ecol Notes **7**: 1002–1004. doi: 10.1111/j.1471-8286.2007.01755.x.

Chenuil A, Le Gac M, Thierry M (2003) Fast isolation of microsatellite loci of very diverse repeat motifs by library enrichment in echinoderm species, *Amphipholis squamata* and *Echinocardium cordatum*. Mol Ecol Notes **3**: 324–327. doi: 10.1046/j.1471-8286.2003.00434.x.

Clement M, Posada D, Crandall KA (2000) TCS: a computer program to estimate gene genealogies. Mol Ecol **9**: 1657–1659. doi: 10.1046/j.1365-294x.2000.01020.x.

Colgan DJ, Byrne M, Rickard E, Castro LR (2005) Limited nucleotide divergence over large spatial scales in the asterinid sea star *Patiriella exigua*. Mar Biol (Berl) **146**: 263–270. doi: 10.1007/s00227-004-1415-6.

Cornuet JM, Luikart G (1996) Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. Genetics **144**: 2001–2014.

Costantini F, Fauvelot C, Abbiati M (2007) Genetic structuring of the temperate gorgonian coral (*Corallium rubrum*) across the western Mediterranean Sea revealed by microsatellites and nuclear sequences. Mol Ecol **16**: 5168–5182.

Diaz-Almela E, Boudry P, Launey S, Bonhomme F, Lapegue S (2004) Reduced female gene flow in the European Xat oyster *Ostrea edulis*. J Hered **95**: 510–516. doi:10.1093/jhered/esh073.

Duran S, Palacin C, Becerro MA, Turon X, Giribet G (2004a) Genetic diversity and population structure of the commercially harvested sea urchin *Paracentrotus lividus* (Echinodermata, Echinoidea). Mol Ecol **13**: 3317–3328. doi: 10.1111/j.1365-294X.2004.02338.x.

Duran S, Pascual M, Estoup A, Turon X (2004b) Strong population structure in the marine sponge *Crambe crambe* (Poecilosclerida) as revealed by microsatellite markers. Mol Ecol **13**: 511–522. doi: 10.1046/j.1365-294X.2004.2080.x.

ElMousadik A, Petit RJ (1996) High level of genetic differentiation for allelic richness among populations of the argan tree *Argania spinosa* (L) Skeels endemic to Morocco. Theor Appl Genet **92**: 832–839. doi: 10.1007/BF00221895.

Estoup A, Rousset F, Michalakis Y, Cornuet JM, Adriamanga M, Guyomard R (1998) Comparative analysis of microsatellite and allozyme markers: a case study investigating microgeographic differentiation in brown trout (*Salmo trutta*). Mol Ecol **7**: 339–353. doi:10.1046/j.1365-294X.1998.00362.x.

Estoup A, Jarne P, Cornuet JM (2002) Homoplasy and mutation model at microsatellite loci and their consequences for population genetics analysis. Mol Ecol **11**: 1591–1604. doi: 10.1046/j.1365-294X.2002.01576.x.

Excoffier L, Laval G, Schneider S (2005) ARLEQUIN (version 3.0): an integrated software package for population genetic data analysis. Bioinformatics Online **1**: 47–50.

Excoffier L, Smouse PE, Quattro JM (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes-application to human mitochondrial-DNA restriction data. Genetics **131**: 479–491.

Féral J-P, Poulin E, Derelle E, Gallardo S, Chambon C (1995) Genetic differentiation of *Echinocardium chordatum* as revealed by allozymes and RNA sequencing. In: Emson R, Smith A, Campbell A (eds) Echinoderm research 1995. Balkema, Rotterdam, pp 41–42.

Geldmacher J, Hoernle K (2000) The 72 Ma geochemical evolution of the Madeira hotspot (eastern North Atlantic): recycling of Palaeozoic (≤500 Ma) oceanic lithosphere. Earth Planet Sci Lett **183**: 73–92. doi: 10.1016/S0012-821X(00)00266-1.

Gerard K, Roby C, Chevalier N, Thomassin B, Chenuil A, Feral JP (2008) Assessment of three mitochondrial loci variability for the crown-of-thorns starfish: a first insight into *Acanthaster* phylogeography. C R Biol **331**: 137–143. doi: 10.1016/j.crvi.2007.11.005.

Goudet J (1995) FSTAT (Version 1.2): a computer program to calculate F-statistics. J Hered **86**: 485–486.

Guo SW, Thompson EA (1992) Performing the exact test of Hardy-Weinberg proportion for multiple alleles. Biometrics **48**: 361–372. doi: 10.2307/2532296.

Hansen MM, Mensberg KLD, Berg S (1999) Postglacial recolonization patterns and genetic relationships among whitefish (*Coregonus* sp.) populations in Denmark, inferred from mitochondrial DNA and microsatellite markers. Mol Ecol **8**: 239–252. doi: 10.1046/j.1365-294X.1999.00557.x.

Harley CDG, Pankey MS, Wares JP, Grosberg RK, Wonham MJ (2006) Color polymorphism and genetic structure in the sea star *Pisaster ochraceus*. Biol Bull **211**: 248–262. doi: 10.2307/4134547.

Harper FM, Addison JA, Hart MW (2007) Introgression versus immigration in hybridizing high-dispersal echinoderms. Evolution **61**: 2410–2418. doi: 10.1111/j.1558-5646.2007.00200.x.

Harper FM, Hart MW (2005) Gamete compatibility and sperm competition affect paternity and hybridization between sympatric Asterias sea stars. Biol Bull **209**: 113–126. doi: 10.2307/3593129.

Hedrick PW (1999) Perspective: highly variable loci and their interpretation in evolution and conservation. Evolution **53**: 313–318. doi: 10.2307/2640768.

Hedrick PW (2005) A standardized genetic differentiation measure. Evolution **59**: 1633–1638.

Hörstadius S (1938) Über die Entwicklung von *Astropecten aranciacus* L. Pubbl Stn Zool Napoli **17**: 221–312.

Hunt A (1993) Effects of contrasting patterns of larval dispersal on the Genetic connectedness of local-populations of 2 intertidal star-fish, Patiriella-Calcar and P-Exigua. Mar Ecol Prog Ser **92**: 179–186. doi: 10.3354/meps092179.

Koehler R (1921) Echinodermes. Faune de France, Lechevalier, Paris, pp 1–210.

Krafsur ES (2002) Population structure of the tsetse fly *Glossina pallidipes* estimated by allozyme, microsatellite and mitochondrial gene diversities. Insect Mol Biol **11**: 37–45. doi: 10.1046/j.0962-1075.2001.00307.x.

Launey S, Ledu C, Boudry P, Bonhomme F, Naciri-Graven Y (2002) Geographic structure in the European flat oyster (*Ostrea edulis* L.) as revealed by microsatellite polymorphism. J Hered **93**: 331–338. doi: 10.1093/jhered/93.5.331.

Lemaire C, Versini JJ, Bonhomme F (2005) Maintenance of genetic differentiation across a transition zone in the sea: discordance between nuclear and cytoplasmic markers. J Evol Biol **18**: 70–80. doi: 10.1111/j.1420-9101.2004.00828.x.

Mariani S, Ketmaier V, de Matthaeis E (2002) Genetic structuring and gene flow in *Cerastoderma glaucum* (Bivalvia : Cardiidae): evidence from allozyme variation at different geographic scales. Mar Biol (Berl) **140**: 687–697. doi: 10.1007/s00227-001-0753-x.

Matsuoka N, Asano H (2003) Genetic variation in northern Japanese populations of the starfish *Asterina pectinifera*. Zool Sci **20**: 985–988. doi: 10.2108/zsj.20.985.

McCartney MA, Keller G, Lessios HA (2000) Dispersal barriers in tropical oceans and speciation in Atlantic and eastern Pacific sea urchins of the genus *Echinometra*. Mol Ecol **9**: 1391–1400. doi: 10.1046/j.1365-294x.2000.01022.x.

McGoldrick DJ, Hedgecock D, English LJ, Baoprasertkul P, Ward RD (2000) The transmission of microsatellite alleles in Australian and North American stocks of the Pacific oyster (*Crassostrea gigas*): selection and null alleles. J Shellfish Res **19**: 779–788.

Meirmans PG (2006) Using the AMOVA framework to estimate a standardized genetic differentiation measure. Evolution **60**: 2399–2402.

Michalakis Y, Excoffier L (1996) A generic estimation of population subdivision using distances between alleles with special reference for microsatellite loci. Genetics **142**:1061–1064.

Mougenot D, Vanney J-R (1982) The Plio-Quarternary sediment drifts of the south Portuguese continental slope. Bull Inst Geologie Bassin Aquitaine **31**: 131–139.

Narum SR (2006) Beyond Bonferroni: less conservative analyses for conservation genetics. Conserv Genet **7**: 783–787. doi: 10.1007/s10592-005-9056-y.

Nei M (1987) Molecular evolutionary genetics. Colombia University Press, New York

Nelson RJ, Cooper G, Garner T, Schnupf P (2002) Polymorphic markers for the sea cucumber *Parastichopus californicus*. Mol Ecol Notes **2**: 233–235. doi: 10.1046/j.1471-8286.2002.00205.x.

Pabst B (1986) Eigenschaften der Dislokation bei drei Seesternarten der Gattung *Astropecten*. Inaugural-Dissertation, Universität Zürich.

Palumbi SR, Grabowsky G, Duda T, Geyer L, Tachino N (1997) Speciation and population genetic structure in tropical Pacific Sea urchins. Evolution **51**: 1506–1517. doi: 10.2307/2411203.

Palumbi SR, Wilson AC (1990) Mitochondrial-DNA diversity in the Sea-Urchins *Strongylocentrotus-Purpuratus* and *Strongylocentrotus-Droebachie*. Evolution **44**: 403–415. doi: 10.2307/2409417.

Parsons TJ, Muniec DS, Sullivan K, Woodyatt N, AllistonGreiner R, Wilson MR, Berry DL, Holland KA, Weedn VW, Gill P, Holland MM (1997) A high observed substitution rate in the human mitochondrial DNA control region. Nat Genet **15**: 363–368. doi: 10.1038/ng0497-363.

Patarnello T, Volckaert F, Castilho R (2007) Pillars of Hercules: is the Atlantic–Mediterranean transition a phylogeographical break? Mol Ecol **16**: 4426–4444. doi: 10.1111/j.1365-294X.2007.03477.x.

Peijnenburg K, Fauvelot C, Breeuwer AJ, Menken SBJ (2006) Spatial and temporal genetic structure of the planktonic *Sagitta setosa* (Chaetognatha) in European seas as revealed by mitochondrial and nuclear DNA markers. Mol Ecol **15**: 3319–3338. doi: 10.1111/ j.1365-294X.2006.03002.x.

Perez-Losada M, Guerra A, Carvalho GR, Sanjuan A, Shaw PW (2002) Extensive population subdivision of the cuttlefish *Sepia offcinalis* (Mollusca : Cephalopoda) around the Iberian Peninsula indicated by microsatellite DNA variation. Heredity **89**: 417–424. doi: 10.1038/sj.hdy.6800160.

Perez-Losada M, Nolte MJ, Crandall KA, Shaw PW (2007) Testing hypotheses of population structuring in the Northeast Atlantic ocean and Mediterranean sea using the common cuttlefish *Sepia officinalis*. Mol Ecol **16**: 2667–2679. doi: 10.1111/j.1365-294X.2007.03333.x.

Perrin C, Roy MS (2000) Rapid and efficient identification of microsatellite loci from the sea urchin, *Evechinus chloroticus*. Mol Ecol **9**: 2221–2223. doi: 10.1046/j.1365-294X.2000.105335.x.

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. Genetics **155**: 945–959.

Quesada H, Zapata C, Alvarez G (1995) A multilocus allozyme discontinuity in the mussel *Mytilus-Galloprovincialis*—the interaction of ecological and life-history factors. Mar Ecol Prog Ser **116**: 99–115. doi: 10.3354/meps116099.

Rogers AR, Harpending H (1992) Population-growth makes waves in the distribution of pairwise genetic-differences. Mol Biol Evol **9**: 552–569.

Roman J, Palumbi SR (2004) A global invader at home: population structure of the green crab, *Carcinus maenas*, in Europe. Mol Ecol **13**: 2891–2898. doi: 10.1111/j.1365-294X.2004.02255.x.

Rousset F (1997) Genetic differentiation and estimation of gene flow from F-statistics under isolation by distance. Genetics **145**: 1219–1228.

Saavedra C, Pena JB (2005) Nucleotide diversity and Pleistocene population expansion in Atlantic and Mediterranean scallops (*Pecten maximus* and *P. jacobaeus*) as revealed by the mitochondrial 16S ribosomal RNA gene. J Exp Mar Biol Ecol **323**: 138–150. doi: 10.1016/j.jembe.2005.03.006.

Shanks AL, Grantham BA, Carr MH (2003) Propagule dispersal distance and the size and spacing of marine reserves. Ecol Appl **13**: S159–S169.

Shaw PW, Pierce GJ, Boyle PR (1999) Subtle population structuring within a highly vagile marine invertebrate, the veined squid *Loligo forbesi*, demonstrated with microsatellite DNA markers. Mol Ecol **8**: 407–417. doi: 10.1046/j.1365-294X.1999.00588.x.

Slatkin M (1995) A measure of population subdivision based on microsatellite allele frequencies. Genetics **139**: 1463.

Slatkin M, Excoffier L (1996) Testing for linkage disequilibrium in genotypic data using the expectation–maximization algorithm. Heredity **76**: 377–383. doi: 10.1038/hdy.1996.55.

Smith MJ, Arndt A, Gorski S, Fajber E (1993) The phylogeny of echinoderm classes based on mitochondrial gene arrangements. J Mol Evol **36**: 545–554. doi: 10.1007/BF00556359.

Stamatis C, Triantafyllidis A, Moutou KA, Mamuris Z (2004) Mitochondrial DNA variation in northeast Atlantic and Mediterranean populations of Norway lobster, *Nephrops norvegicus*. Mol Ecol **13**: 1377–1390. doi: 10.1111/j.1365-294X.2004.02165.x.

Stamatis C, Triantafyllidis A, Moutou KA, Mamuris Z (2006) Allozymic variation in Northeast Atlantic and Mediterranean populations of Norway lobster, *Nephrops norvegicus*. Mark Sci **63**: 875–882. doi: 10.1016/j.icesjms.2006.01.006.

Tajima F (1983) Evolutionary relationship of DNA-sequences in finite populations. Genetics **105**: 437–460.

Tajima F (1989) Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics **123**: 585–595.

Templeton AR, Crandall KA, Sing CF (1992) A cladistic analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping and DNA-sequence data. 3. Cladogram estimation. Genetics **132**: 619–633.

Templeton AR, Routman E, Phillips CA (1995) Separating population structure from population history—a cladistic-analysis of the geographical-distribution of mitochondrial-DNA haplotypes in the Tiger Salamander, *Ambystoma-Tigrinum*. Genetics **140**: 767–782.

Tintore J, Laviolette PE, Blade I, Cruzado A (1988) A study of an intense density front in the Eastern Alboran-Sea–the Almeria-Oran Front. J Phys Oceanogr **18**: 1384–1397.

Tortonese E (1980) Apérçu sommaire sur les asteroidea de la Méditerrannée (histoire, distribution, systematique). Journées d'études sur la systématique évolutive et la biogéographie en Méditerranée, Cagliari, pp 11–19.

Triantafyllidis A, Apostolidis AP, Katsares V, Kelly E, Mercer J, Hughes M, Jorstad K, Tsolou A, Hynes R, Triantaphyllidis C (2005) Mitochondrial DNA variation in the European lobster (*Homarus gammarus*) throughout the range. Mar Biol (Berl) **146**: 223–235. doi: 10.1007/s00227-004-1435-2.

Turan C, Carvalho GR, Mork J (1998) Molecular genetic analysis of Atlanto-Scandian herring (*Clupea harengus*) populations using allozymes and mitochondrial DNA markers. J Mar Biol Assoc UK **78**: 269–283.

Uthicke S, Benzie JAH (2003) Gene flow and population history in high dispersal marine invertebrates: mitochondrial DNA analysis of *Holothuria nobilis* (Echinodermata : Holothuroidea) populations from the Indo-Pacific. Mol Ecol **12**: 2635–2648. doi: 10.1046/j.1365-294X.2003.01954.x.

Van Oosterhout C, Hutchinson WF, Wills DPM, Shipley P (2004) MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. Mol Ecol Notes **4**: 535–538. doi: 10.1111/j.1471-8286.2004.00684.x.

Waters JM, O'Loughlin PM, Roy MS (2004) Cladogenesis in a starfish species complex from southern Australia: evidence for vicariant speciation? Mol Phylogenet Evol **32**: 236–245. doi: 10.1016/j.ympev.2003.11.014.

Waters JM, Roy MS (2004) Phylogeography of a high-dispersal New Zealand sea-star: does upwelling block gene flow? Mol Ecol **13**: 2797–2806. doi: 10.1111/j.1365-294X.2004.02282.x.

Watterson GA (1984) Allele frequencies after a Bottleneck. Theor Popul Biol **26**: 387–407. doi: 10.1016/0040-5809(84)90042-X.

Weir BS, Cockerham CC (1984) Estimating F-statistics for the analysis of population-structure. Evolution Int J Org Evolution **38**: 1358–1370. doi: 10.2307/2408641.

Williams ST (2000) Species boundaries in the starfish genus Linckia. Mar Biol (Berl) **136**: 137–148. doi: 10.1007/s002270050016.

Williams ST, Benzie JAH (1997) Indo-West Pacific patterns of genetic differentiation in the high-dispersal starfish *Linckia laevigata*. Mol Ecol **6**: 559–573. doi: 10.1046/j.1365-294X.1997.00221.x.

Williams ST, Benzie JAH (1998) Evidence of a biogeographic break between populations of a high dispersal starfish: congruent regions within the Indo-West Pacific defined by color morphs, mtDNA, and allozyme data. Evolution **52**: 87–99. doi: 10.2307/2410923.

Williams ST, Jara J, Gomez E, Knowlton N (2002) The marine Indo-West Pacific break: contrasting the resolving power of mitochondrial and nuclear genes. Integr Comp Biol **42**: 941–952. doi: 10.1093/icb/42.5.941.

Xia X, Xie Z (2001) DAMBE: Software package for data analysis in molecular biology and evolution. J Hered **92**: 371–373. doi: 10.1093/jhered/92.4.371.

Yasuda N, Nagai S, Hamaguchi M, Lian CL, Nadaoka K (2006) Development of microsatellite markers for the crown-of-thorns starfish *Acanthaster planci*. Mol Ecol Notes **6**: 141–143. doi: 10.1111/j.1471-8286.2005.01168.x.

Zane L, Ostellari L, Maccatrozzo L, Bargelloni L, Cuzin-Roudy J, Buchholz F, Patarnello T (2000) Genetic differentiation in a pelagic crustacean (*Meganyctiphanes norvegica* : Euphausiacea) from the North East Atlantic and the Mediterranean Sea. Mar Biol (Berl) **136**: 191–199. doi: 10.1007/s002270050676.

Zavodnik D (1960) Echinodermata der Insel Krk. Acta Adriat **9**: 3–19.

Zhan AB, Bao ZM, Lu W, Hu XL, Peng W, Wang ML, Hu JJ (2007) Development and characterization of 45 novel microsatellite markers for sea cucumber (*Apostichopus japonicus*). Mol Ecol Notes **7**: 1345–1348. doi: 10.1111/j.1471-8286.2007.01876.x.

Zulliger D, Ruch M, Tanner S, Ribi G (2008) Characterization of nine microsatellite loci in the sea star *Astropecten aranciacus* and cross-species amplification for related taxa. Mol Ecol Res **8**: 634–636. doi: 10.1111/j.1471-8286.2007.02027.x.

Table 1: *A. aranciacus* intra-population genetic variability at the mitochondrial control region (CR) including flanking regions and four microsatellite loci for nine sampling locations. Indicated are the number of individuals processed (n) for mitochondrial and nuclear markers, as well as the number of total haplotypes (T), the number of location-specific haplotypes found within each sampling location (S) and haplotype (h) and nucleotide (π) diversity. Standard errors are in parenthesis. For microsatellite loci the number of alleles (N_a), expected (H_E) and observed heterozygosity (H_O) and the allelic richness (R_s) are shown. Allelic richness is adjusted for a minimum sample size of 20 diploid individuals. Estimated null allele frequencies (r; van Oosterhout) are shown for populations with significant lower H_O from HWE after false discovery rate (FDR) correction.

	MtDNA					**Microsatellite loci**																
	CR and flanking regions					**Aaran06**					**Aaran09**				**Aaran2/05**				**Aaran2/25**				
Location	n	T	S	h	π	n	N_a	H_O	H_E	R_s	N_a	H_O	H_E	R_s	N_a	H_O	H_E	R_s	N_a	H_O	H_E	R_s	Mean R_s
MAD	19	12	11	0.8363 (0.0199)	0.0029 (0.0004)	21	9	0.810	0.832	8.949	3	0.429	0.547	3.000	6	0.667	0.667	5.952	18	0.700* r = 0.112	0.923	18.000	8.975
FAR	18	15	12	0.9804 (0.0057)	0.0062 (0.0008)	30	11	0.759	0.798	10.149	5	0.724	0.725	4.900	9	0.625	0.710	7.943	22	0.704*** r = 0.121	0.941	18.860	10.463
Total Atlantic	*37*	*27*	*23*	*0.9535 (0.0044)*	*0.0050 (0.0005)*	*51*	*12*	*0.780*	*0.810*	*9.549*	*5*	*0.600*	*0.671*	*3.950*	*9*	*0.640*	*0.695*	*6.948*	*30*	*0.702**** r = *0.127*	*0.951*	*18.430*	*9.719*
LAH	15	13	7	0.9714 (0.0100	0.0069 (0.0010)	22	12	0.864	0.863	11.539	4	0.727	0.718	4.000	11	0.864	0.848	10.545	17	0.727*** r = 0.105	0.943	16.597	10.670
BAN						23	9	0.652°	0.788	8.709	7	0.696	0.620	6.580	10	0.739	0.864	9.810	19	0.619** r = 0.162	0.945	18.567	10.917
MUR	20	20	13	1.0000 (0.0035)	0.0054 (0.0007)	49	14	0.886	0.837	11.627	6	0.818***	0.648	5.115	14	0.773	0.815	10.567	25	0.757* r = 0.096	0.95	19.904	11.803
GAE	20	19	12	0.9947 (0.0040)	0.0072 (0.0009)	28	12	0.893	0.860	10.679	6	0.643	0.604	5.330	11	0.893	0.823	9.773	25	0.857°	0.96	21.138	11.730
Total W. Med.	*55*	*44*	*38*	*0. 9892 (0.0008)*	*0.0064 (0.0005)*	*122*	*15*	*0.838*	*0.838*	*10.639*	*7*	*0.735****	*0.646*	*5.256*	*17*	*0.812*	*0.834*	*10.174*	*36*	*0.750**** r = *0.117*	*0.954*	*19.052*	*11.280*
CRE	20	18	14	0.9895 (0.0043)	0.0065 (0.0008)	21	10	0.750	0.837	10.000	6	0.619°	0.714	5.950	8	0.714	0.816	7.951	24	0.900°	0.968	24.000	11.975
HER	20	19	17	0.9947 (0.0040)	0.0078 (0.0009)	30	12	0.933	0.881	11.197	6	0.467	0.652	5.226	12	0.833°	0.904	11.482	24	0.900	0.958	20.311	12.054
KAV	19	18	14	0.9942 (0.0044)	0.0067 (0.0008)	30	12	0.733	0.847	10.441	5	0.667	0.627	4.333	11	0.897	0.866	9.493	28	0.767*** r = 0.098	0.968	23.131	11.850
Total E. Med.	*59*	*51*	*47*	*0.9930 (0.0007)*	*0.0069 (0.0005)*	*81*	*15*	*0.813*	*0.854*	*10.546*	*8*	*0.580*	*0.654*	*5.170*	*12*	*0.825*	*0.869*	*9.642*	*37*	*0.850**** r = *0.057*	*0.966*	*22.481*	*11.960*
All Pops	*151*	*114*	*100*	*0.9910 (0.0002)*	*0.0064 (0.0003)*	*254*	*17*	*0.818*	*0.838*	*10.598*	*9*	*0.657*	*0.657*	*5.246*	*18*	*0.781*	*0.826*	*9.841*	*46*	*0.774** r = *0.095*	*0.960*	*21.917*	*11.901*

* = $P < 0.05$, ** = $P < 0.01$, *** = $P < 0.001$, ° = values that become non-significant after applying FDR correction.

Table 2: Population pairwise genetic differentiation F_{ST} between sampling locations of *A. aranciacus* based on mtDNA sequence data and F_{ST} and standardized genetic differentiation in four microsatellite loci. Below diagonal: multilocus F_{ST} (Weir & Cockerham 1984) and standardized genetic differentiation (Hedrick 2005) in italic (for microsatellite loci only); above diagonal: *P*-values < 0.05. Significance is indicated in bold after false discovery rate correction.

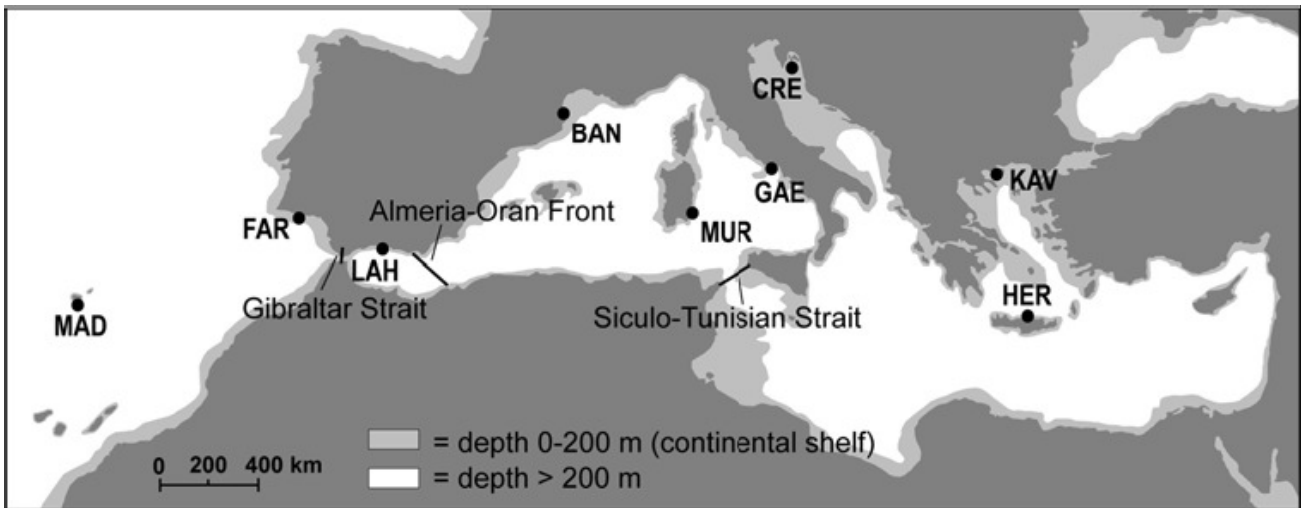
	Atlantic		**West Mediterranean**				**East Mediterranean**		
	MAD	**FAR**	**LAH**	**BAN**	**MUR**	**GAE**	**CRE**	**HER**	**KAV**
mtDNA									
MAD	-	**0.0000**	**0.0000**		**0.0000**	**0.0000**	**0.0000**	**0.0000**	**0.0000**
FAR	0.1796	-						0.0488	0.0396
LAH	0.1666	-0.0253	-						
BAN				-					
MUR	0.1711	0.0039	-0.0045		-				0.0397
GAE	0.1884	-0.0014	-0.0151		-0.0147	-		0.0489	0.0381
CRE	0.1378	-0.0043	0.0011		0.0020	0.0105	-		
HER	0.1630	0.0428	0.0257		0.0337	0.0437	0.0026	-	
KAV	0.1663	0.0463	0.0351		0.0419	0.0510	0.0007	-0.0317	-
microsatellite loci									
MAD	-	**0.0046**	**0.0040**	**0.0039**	**0.0114**	**0.0160**	**0.0315**	**0.0013**	**0.0029**
FAR	0.0254 *0.1079*	-		**0.0059**		**0.0273**		**0.0030**	**0.0083**
LAH	0.0299 *0.1443*	-0.0010 *0.0000*	-						
BAN	0.0312 *0.1361*	0.0231 *0.1129*	0.0052 *0.0296*	-					
MUR	0.0165 *0.0716*	0.0010 *0.0048*	-0.0027 *0.0000*	0.0016 *0.0080*	-				
GAE	0.0178 *0.0809*	0.0125 *0.0635*	0.0068 *0.0400*	0.0021 *0.0111*	-0.0039 *0.0000*	-			
CRE	0.0197 *0.0914*	0.0073 *0.0383*	0.0032 *0.0198*	0.0094 *0.0514*	0.0024 *0.0129*	-0.0004 *0.0000*	-		
HER	0.0305 *0.1529*	0.0213 *0.1204*	0.0055 *0.0366*	0.0051 *0.0302*	0.0061 *0.0366*	0.0034 *0.0207*	0.0017 *0.0110*	-	
KAV	0.0251 *0.1177*	0.0167 *0.0879*	-0.0003 *0.0000*	0.0003 *0.0016*	0.0020 *0.0108*	-0.0009 *0.0000*	0.0004 *0.0024*	-0.0069 *0.0000*	-

Table 3: Analysis of molecular variance (AMOVA) based on pairwise distances in *A. aranciacus* for different groupings of Atlantic (A) versus Mediterranean (M) and within the Mediterranean for eastern (EM) versus western (WM) samples. Grouping the sample LAH to a basin other than the western Mediterranean is indicated by +, removing it by -. The table includes the degrees of freedom (df) for both mitochondrial DNA (mtDNA) and microsatellite loci (msats) and the F-statistics. Only *P*-values > 0.05 are listed, and significance after false discovery rate correction is indicated in bold.

	df		**Percentage of variation**		**Fixation indices**			***P***	
Grouping	**mtDNA**	**msats**	**mtDNA**	**msats**		**mtDNA**	**msats**	**mtDNA**	**msats**
(A) A+ vs. M									
Among groups	1	1	1.39	0.52	F_{CT} =	0.013	0.005	-	**0.035**
Within groups	6	7	3.94	0.500	F_{SC} =	0.039	0.005	**0.005**	0.035
Within populations	143	487	94.68	98.98	F_{ST} =	0.053	0.010	**0.000**	**0.002**
(B) A vs. M									
Among groups	1	1	2.11	0.80	F_{CT} =	0.021	0.008	-	**0.029**
Within groups	6	7	3.73	0.45	F_{SC} =	0.038	0.005	**0.000**	0.049
Within populations	143	487	94.15	98.75	F_{ST} =	0.058	0.013	**0.000**	**0.002**
(C) A+ vs. WM vs. EM									
Among groups	2	2	2.47	0.48	F_{CT} =	0.024	0.005	0.043	**0.014**
Within groups	5	6	2.79	0.39	F_{SC} =	0.028	0.004	**0.029**	-
Within populations	143	487	94.73	99.13	F_{ST} =	0.052	0.009	**0.000**	**0.002**
(D) A vs. WM vs. EM									
Among groups	2	2	3.24	0.60	F_{CT} =	0.032	0.006	0.014	**0.007**
Within groups	5	6	2.22	0.32	F_{SC} =	0.022	0.003	**0.002**	-
Within populations	143	487	94.54	99.08	F_{ST} =	0.054	0.009	**0.000**	**0.002**
(E) WM vs. EM									
Among groups	1	1	3.79	0.37	F_{CT} =	0.037	0.004	-	**0.028**
Within groups	4	5	-1.01	-0.06	F_{SC} =	-0.010	-0.001	-	-
Within populations	108	389	97.21	99.69	F_{ST} =	0.027	0.003	-	-
(F) WM- vs. EM									
Among groups	1	1	3.94	0.47	F_{CT} =	0.039	0.005	-	-
Within groups	3	4	-0.96	-0.13	F_{SC} =	-0.010	-0.001	-	-
Within populations	94	346	97.02	99.66	F_{ST} =	0.029	0.003	-	-

- = non significant ($P >= 0.05$)

Figure 1: Map showing the sampling sites of *A. aranciacus*.



MAD Madeira, *FAR* Faro, *LAH* La Herradura, *BAN* Banyuls, *MUR* Muravera, *GAE* Gaeta, *CRE* Cres, *HER* Heraklion, *KAV* Kavala

Figure 2: Parsimony network showing relationships among all *Astropecten aranciacus* mtDNA haplotypes. Haplotypes are colored according to the region: Atlantic (*gray*), western Mediterranean (*black*), and eastern Mediterranean (*white*). *Circle* size is proportional to the number of samples with the same haplotype and size of pie slices to the number of samples from the same region. *Each line* in the network represents a single mutational change. *Small dots* indicate intermediate haplotypes not observed in the sample but necessary to link all observed haplotypes to the network

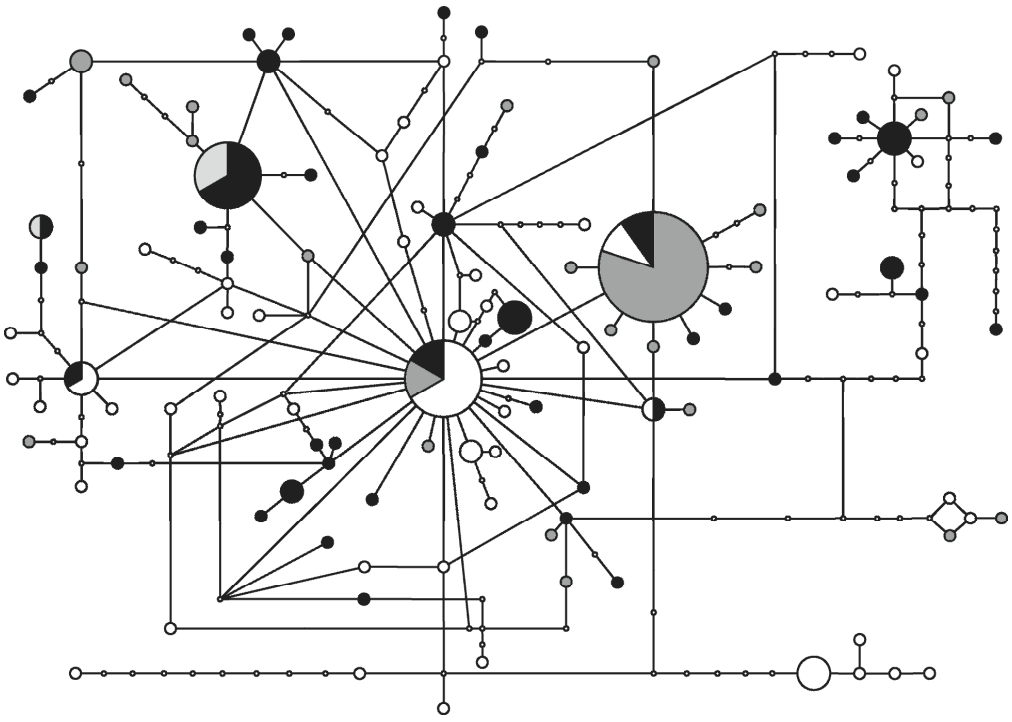


Figure 3: Results of the structure analysis using microsatellite data indicating the average ¡log PR(*X*|*K*) estimate with standard error bars for each number of populations (*K*)

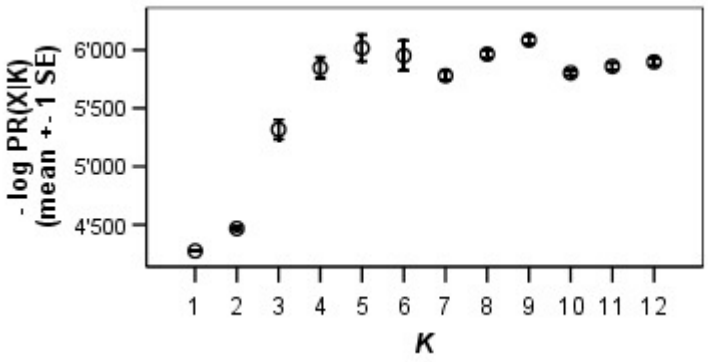
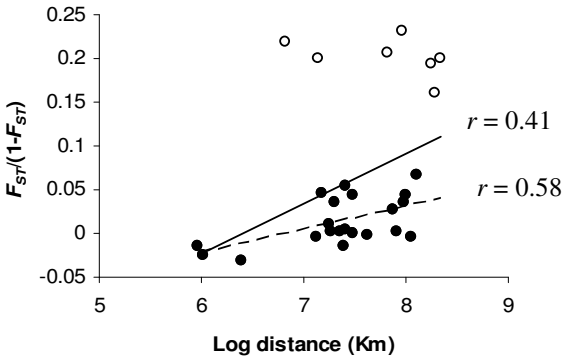
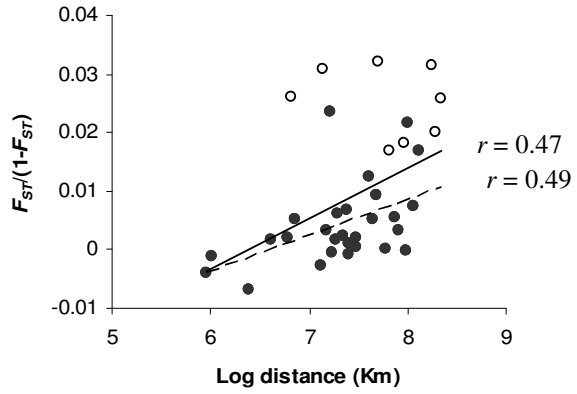


Figure 4: Genetic differentiation (computed as $F_{ST}/(1-F_{ST})$) versus the logarithm of the geographic distance in km in *A. aranciacus* based on a) mtDNA sequence data and b) microsatellite data. Full lines represent the regression line over all pairwise F_{ST} , dotted lines represent the regression line omitting the location MAD.

a) mtDNA



b) microsatellite loci



○ = pairs with MAD ● = pairs without MAD

Appendix I: Absolute haplotype frequencies of all sampled locations of *Astropecten aranciacus*. The column 'corresponding samples' indicates the specimens' identifier.

Haplotype	MAD	FAR	LAH	MUR	GAE	CRE	HER	KAV	corresponding samples
1	8		1			1			ma1, ma2, ma6, ma7, ma11, ma16, ma17, ma21, cr15, sp9
2	1								ma3
3	1								ma4
4	1								ma5
5	1								ma10
6	1								ma12
7	1								ma13
8	1								ma14
9	1								ma15
10	1								ma18
11	1								ma19
12	1								ma20
13		1							po18
14		1							po22
15		1							po23
16		1							po24
17		1							po27
18		1							po28
19		1							po30
20		1	1						po34, sp4
21		2							po35, po44
22		1							po38
23		1							po39
24		1							po45
25		1							po48
26			1						sp1
27		2	3		1				sp2, sp3, sp14, po26, po42, na6
28			1						sp5
29			1						sp10
30			1						sp12
31			1						sp15
32			1						sp16
33			1						sp17
34				1					mu71
35				1					mu72
36				1					mu73
37			1	1					mu75, sp11
38				1					mu80
39				1	1				mu81, na2
40				1	1				mu82, na24
41				1					mu89
42				1					mu90
43				1					mu91
44				1					mu109
45				1					mu111
46			1	1	1				mu112, na21, s18
47				1	1				mu117, na4
48				1					mu120
49				1					mu121
50				1					mu122
51				1					mu123
52					1				na3
53					1				na7
54					1				na8
55					1				na9
56					1				na12
57					1				na13
58					1				na14
59					1				na17
60					1				na18
61					1				na19
62					1				na20

Haplotype	MAD	FAR	LAH	MUR	GAE	CRE	HER	KAV	corresponding samples
63				1	2				na25, na27, mu119
64					1				na28
65						1			cr16
66						1			cr17
67						1			cr183
68						2			cr184, cr195
69						1			cr185
70						1			cr187
71						1			cr188
72		2		1		2		2	cr189, cr201, ka2, ka28, mu43, po4, po5
73						1			cr190
74						1			cr193
75					1	1		1	cr194, ka9, na22
76						1			cr196
77						1			cr198
78						1	1		cr199, kr5
79						1			cr200
80						1			cr202
81						1			cr203
82							2	1	kr1, kr2, ka12
83							1		kr3
84							1		kr6
85							1		kr7
86							1		kr8
87							1		kr9
88							1		kr10
89							1		kr13
90							1		kr15
91							1		kr16
92							1		kr19
93							1		kr22
94							1		kr23
95							1		kr24
96							1		kr26
97							1		kr27
98							1		kr29
99							1		kr30
100								1	ka1
101								1	ka3
102								1	ka4
103								1	ka5
104								1	ka7
105								1	ka8
106								1	ka11
107			1					1	ka14, sp19
108								1	ka15
109								1	ka16
110								1	ka18
111								1	ka22
112								1	ka23
113								1	ka26
114								1	ka29

Appendix II: Allele-frequencies at four microsatellite loci for *A. aranciacus* from nine sampling locations. Number of individuals sampled is indicated by *n*.

		Location										
Locus	**Allele**	**MAD** ***n*** **=21**	**FAR** ***n*** **= 38**	**LAH** ***n*** **= 22**	**BAN** ***n*** **= 23**	**MUR** ***n*** **= 44**	**GAE** ***n*** **= 28**	**CRE** ***n*** **= 21**	**HER** ***n*** **= 30**	**KAV** ***n*** **= 30**	**all weighted**	**all un-weighted**
Aaran06												
	84	0.048	0.034	0.023	0	0.034	0	0.025	0.033	0.033	0.026	0.026
	86	0.333	0.414	0.295	0.391	0.352	0.25	0.325	0.317	0.25	0.326	0.325
	88	0.071	0.069	0.136	0.109	0.102	0.089	0.1	0.133	0.15	0.107	0.107
	90	0.167	0.121	0.159	0.217	0.136	0.196	0.2	0.117	0.1	0.152	0.157
	92	0.048	0.034	0.023	0.043	0.068	0.018	0	0	0.033	0.032	0.03
	95	0	0.017	0.023	0	0.011	0	0.025	0	0	0.008	0.008
	97	0	0	0	0	0	0.018	0	0	0	0.002	0.002
	99	0.167	0.052	0.068	0.087	0.057	0.054	0.05	0.067	0.033	0.067	0.07
	101	0.095	0.086	0.045	0	0.057	0.036	0	0.033	0.05	0.047	0.045
	103	0	0.017	0	0.043	0.057	0.018	0.075	0.017	0.083	0.036	0.034
	105	0.048	0.086	0.091	0.022	0.045	0.179	0.125	0.117	0.133	0.093	0.094
	107	0.024	0.069	0.091	0.065	0.023	0.089	0.025	0.117	0.083	0.065	0.065
	109	0	0	0.023	0	0.023	0.018	0.05	0.017	0.017	0.016	0.016
	111	0	0	0	0	0	0	0	0	0.033	0.004	0.004
	113	0	0	0	0	0.011	0.036	0	0.017	0	0.008	0.007
	117	0	0	0	0	0	0	0	0.017	0	0.002	0.002
	122	0	0	0.023	0.022	0.023	0	0	0	0	0.008	0.007
Aaran09												
	115	0	0	0	0.022	0	0.036	0.048	0	0.033	0.014	0.015
	117	0.119	0.207	0.205	0.043	0.068	0.054	0.238	0.167	0.167	0.135	0.141
	119	0.619	0.328	0.364	0.543	0.466	0.536	0.452	0.533	0.517	0.482	0.484
	121	0	0	0	0.022	0.011	0.018	0	0	0.017	0.008	0.008
	122	0	0.069	0.091	0.043	0.068	0	0	0	0	0.032	0.03
	123	0.262	0.362	0.341	0.304	0.364	0.339	0.19	0.267	0.25	0.304	0.298
	124	0	0	0	0	0	0	0.048	0	0	0.004	0.005
	125	0	0	0	0	0	0	0	0.017	0	0.002	0.002
	128	0	0.034	0	0.022	0.023	0.018	0.024	0.017	0.017	0.018	0.017
Aaran2/05												
	193	0.024	0.017	0	0.043	0.034	0.018	0.095	0.034	0.083	0.038	0.039
	194	0	0	0	0	0.011	0	0	0	0	0.002	0.001
	195	0.548	0.5	0.318	0.196	0.364	0.357	0.357	0.259	0.2	0.342	0.344
	196	0	0	0	0	0	0	0	0	0.017	0.002	0.002
	197	0	0.017	0.023	0.043	0.057	0.107	0.024	0.069	0.067	0.049	0.045
	199	0.071	0.052	0.091	0.043	0.023	0.018	0	0.034	0.1	0.047	0.048
	201	0	0.017	0.023	0.043	0.034	0.089	0.095	0.138	0.067	0.057	0.056
	203	0.143	0.034	0.091	0.13	0.057	0.071	0.095	0.086	0.1	0.085	0.09
	204	0	0	0	0	0.011	0	0	0	0.033	0.006	0.005
	205	0.095	0.138	0.114	0.174	0.182	0.161	0.19	0.103	0.067	0.138	0.136
	206	0.119	0.069	0.114	0.065	0.091	0.036	0	0.086	0.067	0.073	0.072
	207	0	0.155	0.159	0.239	0.102	0.107	0.095	0.172	0.15	0.132	0.131
	208	0	0	0.023	0	0	0	0	0	0	0.002	0.003
	209	0	0	0	0	0	0.018	0.048	0.017	0.05	0.014	0.015
	210	0	0	0.023	0.022	0	0	0	0	0	0.004	0.005
	211	0	0	0.023	0	0.011	0.018	0	0	0	0.006	0.006
	212	0	0	0	0	0.011	0	0	0	0	0.002	0.001
	213	0	0	0	0	0.011	0	0	0	0	0.002	0.001

		Location										
Locus	**Allele**	**MAD** ***n*** **=21**	**FAR** ***n*** **= 38**	**LAH** ***n*** **= 22**	**BAN** ***n*** **= 23**	**MUR** ***n*** **= 44**	**GAE** ***n*** **= 28**	**CRE** ***n*** **= 21**	**HER** ***n*** **= 30**	**KAV** ***n*** **= 30**	**all weighted**	**all un-weighted**
Aaran2/25												
	70	0	0.111	0.091	0	0.054	0.089	0.075	0.033	0.033	0.055	0.054
	72	0	0	0	0	0	0	0	0.017	0	0.002	0.002
	74	0	0.056	0.045	0.119	0.068	0.071	0.025	0.083	0.117	0.068	0.065
	76	0	0.037	0.023	0	0.054	0	0	0.05	0.05	0.028	0.024
	80	0.225	0	0	0.071	0	0.018	0.05	0	0	0.032	0.04
	83	0	0.037	0.159	0.048	0.027	0	0	0.05	0.017	0.036	0.037
	85	0.025	0	0.045	0	0.027	0	0	0.05	0	0.017	0.016
	87	0.025	0.093	0.091	0.143	0.081	0.071	0.025	0.067	0.033	0.07	0.07
	89	0	0	0	0	0	0.018	0	0	0	0.002	0.002
	93	0	0	0	0.024	0	0	0	0	0	0.002	0.003
	95	0	0	0	0.048	0	0.018	0	0	0	0.006	0.007
	96	0	0	0	0	0	0.018	0	0.017	0.033	0.009	0.008
	97	0	0	0.045	0	0.014	0.018	0	0	0	0.009	0.009
	98	0	0	0	0	0	0.018	0	0.05	0.083	0.019	0.017
	100	0.05	0.037	0.045	0.024	0.027	0.071	0.05	0.017	0.067	0.043	0.043
	102	0.025	0	0	0.095	0.027	0.071	0.05	0.033	0	0.032	0.034
	104	0.1	0.056	0.023	0.024	0.027	0.089	0.025	0.083	0.067	0.055	0.055
	106	0.05	0.019	0.045	0	0.068	0	0.075	0.067	0.05	0.043	0.041
	108	0.025	0.019	0	0	0	0.018	0	0	0	0.006	0.007
	110	0.025	0.019	0.068	0.024	0.041	0	0.025	0	0.017	0.023	0.024
	111	0.075	0	0	0	0	0	0	0	0	0.006	0.008
	113	0	0	0	0	0	0.018	0.05	0	0.017	0.009	0.009
	115	0	0	0	0	0.014	0.018	0.025	0.05	0	0.013	0.012
	117	0.025	0.019	0	0	0.027	0.018	0.025	0.033	0.017	0.019	0.018
	119	0.1	0.074	0	0.024	0.027	0	0	0.017	0.067	0.034	0.034
	121	0.025	0	0	0	0.014	0.036	0	0.033	0	0.013	0.012
	123	0	0.148	0.068	0.071	0.081	0.054	0.125	0.017	0.067	0.07	0.07
	125	0	0	0.091	0.048	0.027	0	0.025	0.05	0	0.026	0.027
	127	0	0	0	0	0.014	0.018	0	0	0	0.004	0.003
	129	0	0.019	0.068	0.048	0.027	0.036	0.025	0.017	0.017	0.028	0.028
	131	0.025	0	0	0	0	0	0	0	0.017	0.004	0.005
	133	0	0.019	0	0	0	0	0	0.017	0	0.004	0.004
	135	0.025	0	0	0	0	0	0.05	0.017	0	0.009	0.01
	137	0	0	0	0.024	0.014	0.036	0.025	0	0.017	0.013	0.013
	139	0.05	0.037	0	0.024	0.027	0.036	0.025	0.033	0.033	0.03	0.029
	141	0.1	0.111	0.023	0.095	0.149	0.089	0.075	0	0.05	0.079	0.077
	143	0	0.019	0.045	0.024	0.054	0.036	0.025	0	0.017	0.026	0.024
	145	0	0	0	0	0.014	0	0.025	0.017	0.067	0.015	0.014
	147	0	0	0.023	0	0	0.018	0.025	0.017	0.017	0.011	0.011
	149	0	0.019	0	0	0	0	0.025	0	0	0.004	0.005
	151	0	0.019	0	0	0	0	0	0.017	0	0.004	0.004
	153	0.025	0	0	0	0	0	0	0.033	0.033	0.011	0.01
	155	0	0	0	0	0	0	0	0.017	0	0.002	0.002
	159	0	0.019	0	0	0	0	0	0	0	0.002	0.002
	161	0	0.019	0	0.024	0	0	0	0	0	0.004	0.005
	163	0	0	0	0	0	0	0.05	0	0	0.004	0.006

GENERAL DISCUSSION AND CONCLUSIONS

This thesis addressed questions related to speciation patterns and genetic diversity in a species-rich genus of sea stars (*Astropecten*) with planktotrophic larvae. In order to obtain a comprehensive picture, these questions were investigated at both an inter- and intra-species level. In addition, an attempt was made to reconstruct the evolutionary history of this genus.

The molecular phylogeny reconstructed in Chapter I using mitochondrial DNA (mtDNA) sequence data shows the relationships of species on a global scale. Phylogenetic inference grouped most species by their geographic region and resulted in three major clades: 1. Indo-West Pacific; 2. American west and east coast; and 3. East Atlantic and Mediterranean. An exception to this phylogenetic pattern is *A. aranciacus*, which, according to the results from the molecular dating (Chapter II), belongs to an old lineage dating back to the Mid-Miocene, as it does not show a clear phylogenetic affinity to species of a particular region. Deep sea lineages in the South Pacific have also split off early, and therefore the phylogenetic grouping with other species from that region is not well supported either. It would be interesting to include more deep sea specimens from other regions such as the Atlantic, the American region and the Indian Ocean, and to investigate whether they also belong to old lineages without a clear phylogenetic affinity to a particular geographic region. *A. irregularis* from the deep sea in the East Atlantic (Portugal), however, does not belong to a particularly ancient lineage and clearly groups with other species of the Atlanto-Mediterranean area. Although this could suggest that deep sea species are not generally ancient lineages, further investigations are necessary to obtain more clarity.

Furthermore, the present study showed that species with previously assumed wide distribution areas throughout the Indian Ocean and West Pacific, such as *A. polyacanthus* and *A. indicus*, in fact exhibit genetically distinct variations and can therefore be considered as species-complexes. It is also possible that these species-complexes include cryptic species. Similarly, morphological variations in *A. irregularis*, which is widely distributed throughout the East Atlantic and Mediterranean region, are also expressed genetically and can be raised to the species-level. This indicates that the dispersal capacity of the planktotrophic larval stage in these species may not be as high as presumed.

In order to estimate the larval dispersal in *Astropecten*, the population genetic structure of *A. aranciacus* was assessed across the Atlanto-Mediterranean region using mitochondrial DNA (mtDNA) and microsatellite loci (Chapter IV). An isolation-by-distance pattern was revealed in this species, which can have a planktotrophic larval stage lasting up to 60 days (Hörstadius 1938), indicating that there is a slight genetic differentiation between populations over the investigated area. This genetic pattern has also been revealed in some other high dispersal marine invertebrates, such as in the European flat oyster *Ostrea edulis* using microsatellite loci and mtDNA sequence data (Launey et al. 2002; Diaz-Almela et al. 2004) and in the

pelagic crustacean *Megenyctiphanes norvegica* (Zane et al. 2000) using mtDNA data only. It is not clear though whether the differentiation in *Astropecten* is enforced by potential marine barriers, such as the Strait of Gibraltar or the Siculo-Tunesian Strait. Genetic differentiation between populations (F_{ST}), however, was mostly insignificant, with the exception of the population from the Island of Madeira in the Atlantic to other populations along the continental coast line, indicating that gene flow to this Island is somewhat restricted. This is either caused by the relative remoteness of this Island or could be due to the prevailing current pattern separating this population from the populations along the continental coast line.

Although gene flow in *A. aranciacus* is apparently not significantly blocked by potential marine barriers, such as the Strait of Gibraltar or the Siculo-Tunisian Strait, this example might not necessarily reflect the genetic structure of other *Astropecten* species with shorter larval stages and an overlapping distribution range. The low population structure in *A. aranciacus* does not reveal any possible speciation processes in this area and is somewhat in contradiction to the high species diversity found in this genus. It would therefore be useful to assess the population genetic structure of a co-occurring species in order to determine whether this pattern is common among *Astropecten*. *A. irregularis pentacanthus* would be a suitable taxon, as it is the only other known species with a distribution range similar to *A. aranciacus*. Similarly, parallel assessments of the genetic structure in species from other areas, such as the American region, the Indian Ocean and the West Pacific would help to determine if a comparable isolation-by-distance pattern can be found in species of other geographic areas and which factors might lead to speciation.

In the species-complexes *A. indicus* and *A. polyacanthus* of the Indo-West Pacific, sibling species occur allopatrically and most likely have speciated by vicariance. However, it is noteworthy that most extant sister taxa occur sympatrically, which might have occurred by allopatric speciation due to vicariance with subsequent recolonization of the geminate species. This is a likely scenario for e.g., Mediterranean species which have originated during the Pleistocene known for its repeated glaciation periods. On the other hand, speciation might also have occurred sympatrically in other regions. The possibility of sympatric speciation has also been considered for sea urchins, such as *Diadema* (Lessios et al. 2001) and *Lytechinus* (Zigler and Lessios 2004), and although it has been rejected in *Diadema,* it is not clear whether all *Lytechinus* species are the result of allopatric speciation. However, reliable methods to determine sympatric speciation have not yet been developed, and the occurrence of sympatric speciation *per se* is highly controversial (Mayr 1982; Berlocher and Feder 2002; Bolnick and Fitzpatrick 2007). Conceivably, this hypothesis could be tested by studying, e.g., the reproductive biology of closely related sympatric species. It appears that hybridization occurs in *Astropecten*, as we have encountered intermediate forms of *A. platyacanthus* and *A. bispinosus* in the Mediterranean Sea, indicating that reproductive isolation between these two geminate species is not completely finalized.

Although morphology suggested several close relationships between species in geographically distant areas, molecular data showed evidence of a clear phylogenetic separation of these regions, indicating that

morphological convergence has taken place in *Astropecten*. This is in line with the results obtained from the molecular dating (Chapter II), which suggest that the speciation rate in this genus has been high and that the diversity within extant species has emerged in a rather short time. Also, opposed to many other marine invertebrates, no extant sister taxa pair across the Isthmus of Panama could be identified in *Astropecten*. Species-rich groups, such as *Astropecten*, may be less likely to contain true geminates separated by the final rise of the Isthmus, because high speciation rates allow groups of closely related species to evolve from the initial geminates on either side. Contrarily, groups containing only few species with wide transoceanic ranges, where profound marine barriers provide one of the few opportunities to speciate, are therefore more likely to harbour trans-isthmian geminate pairs (Collin 2003). On the other hand, true geminates of supposed trans-isthmian sister species may have been removed due to extinction (Marko and Jackson 2001). Nevertheless, studies on some species-rich groups of marine invertebrates, such as gastropods of the genera *Conus* (Duda and Kohn 2005) and *Echinolittorina* (Williams and Reid 2004), as well as crustaceans belonging to the genera *Alpheus* (Anker et al. 2008a; Anker et al. 2008b) and *Petrolisthes* (Hiller et al. 2006), provided evidence of sister species divided by the rise of the Isthmus, thus demonstrating the uniqueness of evolutionary history in each taxon.

Many difficulties were encountered in this study when identifying species based on morphological species descriptions, and it became evident that a taxonomic revision of this genus is urgent. As shown by the present study, molecular markers can provide a valuable complement to our predecessors' morphological work and help to clarify unresolved issues of evolutionary history and systematics.

In a few cases, genetic data support the view that some described species should be synonymized. In the Indo-Pacific region, molecular data revealed a few close relationships between supposedly conspecific populations of *Astropecten* from geographically distant locations such as between *A. polyacanthus* specimens from Hawaii and from Dubai. Waters et al. (2004) suggested that large distance distribution throughout the Indo-Pacific occurs even in some asterinids lacking planktonic larval stages. To what extent ocean currents or other mechanisms have lead to this pattern in Indo-Pacific *Astropecten* remains to be determined.

Finally, even though the genus *Astropecten* is known for its species-richness, the diversity in this genus might yet be underestimated, as many new species, e.g., from several deep-sea populations discovered in the South Pacific, remain to be described.

References

Anker A, Hurt C, Knowlton N (2008a) Revision of the *Alpheus formosus* Gibbes, 1850 complex, with redescription of *A-formosus* and description of a new species from the tropical western Atlantic (Crustacea : Decapoda : Alpheidae). Zootaxa **1707**: 1-22.

Anker A, Hurt C, Knowlton N (2008b) Revision of the *Alpheus websteri* Kingsley, 1880 species complex (Crustacea : Decapoda : Alpheidae), with revalidation of *A-arenensis* (Chace, 1937). Zootaxa **1694**: 51-68.

Berlocher SH, Feder JL (2002) Sympatric speciation in phytophygous insects: moving beyond controversy? Annu Rev Entomol **46**: 773-815.

Bolnick DI, Fitzpatrick BM (2007) Sympatric speciation: Models and empirical evidence. Annu Rev Ecol Evol Syst **38**: 459-487.

Collin R (2003) Phylogenetic relationships among calyptraeid gastropods and their implications for the biogeography of marine speciation. Systematic Biology **52**: 618-640.

Diaz-Almela E, Boudry P, Launey S, Bonhomme F, Lapegue S (2004) Reduced female gene flow in the European Xat oyster *Ostrea edulis*. J Hered **95**:510–516. doi: 10.1093/jhered/esh073.

Duda TF, Kohn AJ (2005) Species-level phylogeography and evolutionary history of the hyperdiverse marine gastropod genus *Conus*. Molecular Phylogenetics and Evolution **34**: 257-272.

Hiller A, Kraus H, Almon M, Werding B (2006) The *Petrolisthes galathinus* complex: Species boundaries based on color pattern, morphology and molecules, and evolutionary interrelationships between this complex and other Porcellanidae (Crustacea : Decapoda : Anomura). Mol Phylogenet Evol **40**: 547-569.

Hörstadius S (1938) Über die Entwicklung von *Astropecten aranciacus* L. Pubbl Stn Zool Napoli **17**: 221-312.

Launey S, Ledu C, Boudry P, Bonhomme F, Naciri-Graven Y (2002) Geographic structure in the European flat oyster (*Ostrea edulis* L.) as revealed by microsatellite polymorphism. J Hered **93**:331–338. doi:10.1093/jhered/93.5.331.

Lessios HA, Kessing BD, Pearse JS (2001) Population structure and speciation in tropical seas: Global phylogeography of the sea urchin Diadema. Evolution **55**: 955-975.

Marko PB, Jackson JBC (2001) Patterns of morphological diversity among and within arcid bivalve species pairs separated by the Isthmus of Panama. J. Paleontol. **75**: 590-606.

Mayr E (1982) The growth of biological thought: Diversity, evolution and inheritance. Belknap Press, Cambridge, MA, pp 974.

Waters JM, O'Loughlin PM, Roy MS (2004) Molecular systematics of some Indo-Pacific asterinids (Echinodermata, Asteroidea): does taxonomy reflect phylogeny? Mol Phylogenet Evol **30**: 872-878. doi: 10.1016/j.ympev.2003.08.019.

Williams ST, Reid DG (2004) Speciation and diversity on tropical rocky shores: A global phylogeny of snails of the genus *Echinolittorina*. Evolution **58**: 2227-2251.

Zane L, Ostellari L, Maccatrozzo L, Bargelloni L, Cuzin-Roudy J, Buchholz F, Patarnello T (2000) Genetic differentiation in a pelagic crustacean (*Meganyctiphanes norvegica* : Euphausiacea) from the North East Atlantic and the Mediterranean Sea. Mar Biol (Berl) **136**:191–199. doi:10.1007/s002270050676.

Zigler KS, Lessios HA (2004) Speciation on the coasts of the new world: Phylogeography and the evolution of bindin in the sea urchin genus *Lytechinu*s. Evolution **58**: 1225-1241.

ACKNOWLEDGEMENTS

I am profoundly grateful for the supervision and guidance I received from **Georg Ribi** while planning and carrying out my thesis. Without his willingness to accept me as a Ph.D. student and his assistance while designing the framework for my studies, this thesis would not have been possible. Georg was also very supportive when applying for grants and collected several specimens which I used in this study. Most importantly, he always provided a pleasant and "stress-free" environment to work in.

The late **Paul Ward** I would like to thank for being on my Ph.D. committee and for his trust and assistance in many ways. He was also very supportive in providing a working space for me and financial means for the lab.

Furthermore, I thank the more recent members of my Ph.D. committee, **Wolf Blankenhorn** and **Tony Wilson**, for their willingness to take on this project at a later stage, for their support in respect to several analytical and statistical issues and for critically reviewing the manuscripts.

I am greatly indebted to **Harilaos Lessios**, who enabled and hosted my visits to the Smithsonian Tropical Research Institute in Panama, where I was able to collect specimens, produce sequence data and learn a lot about phylogenetics. Thanks to him and the people working in his lab, in particular **Axel** and **Ligia Calderón, Renate Sponer** and **Alexandra Hiller**, I experienced two enriching visits to the tropics. Haris was like an additional supervisor to me, and his valuable input and experience were very essential to the success of this thesis.

Samuel Tanner did his Master thesis as part of one of my studies, and I would like to thank him for his exact working style, his assistance in the lab and his endurance with publishing the manuscript. It was great to work with him and also to discuss and have chats with him about a lot of other things.

Another initiator of this thesis was **Markus Ruch**. I thank him for the load of lab work and conceptual thinking he did as an incentive for this thesis.

Marco Bernasconi, **Peter Wandeler**, **Thomas Bucher**, **Andy Pemberton, Yves Choffat** and **Marco Demont** were always helpful with technical advice in lab related and computer issues, and I would like to thank all of them for their assistance. I particularly appreciated the availability and help of Marco B. with questions related to phylogenetics and the input on population genetic matters from Peter.

My visit to the Hawaiian Institute of Marine Biology was enabled by **Rob Toonen**, who was also my supervisor during my stay and was very helpful in several ways. Many thanks go to him for hosting my visit and for his valuable input.

Rosemarie Keller, **Ursula Briegel** and **Sabine Marty** deserve my gratitude for taking care of all administrative things and **Lukas Keller** for his support and understanding for my extra-universitary situation.

With my lab colleague and friend **Sonja Sbilordo** I shared many years as a "fellow sufferer". I highly appreciated her presence and open ear for all my little work and non-work related aches and pains and for her indispensable moral support.

I also thank **Markus Huber** for being such an entertaining office mate and for reminding me to do the reference list. I am aware that as a "non-bivalve" person I was very privileged to have all these interesting taxonomical discussions with him.

The understanding from my former boss **Martin Hägeli** and my present boss **Urs-Peter Wepfer** enabled me to arrange an ideal set up for financing my Ph.D., managing my family life and going abroad for my thesis longer than could be taken for granted. I thank them both for their exceptional flexibility.

For the support of my loving **mother**, **father** and **brother** I am deeply grateful. Their help was present in countless ways and has significantly contributed to the success of this thesis.

A special thank goes to my friend **Christoph Glattfelder**. In spite of many difficult moments during this thesis he was at my side and was very supportive and understanding. He endured all my ups and downs and also directly contributed to this thesis by collecting samples and accompanying me on dives.

Most of all I thank my husband **Paulo Morais** for his love and support in all possible ways. I am profoundly grateful for his willingness to follow me around the globe, for critically reading the manuscripts and providing useful input, for morally supporting me and for the numerous hours of babysitting. His assistance and contribution to my wellbeing were at the foundation of completing this thesis.

CURRICULUM VITAE

Name	Zulliger
Vorname	Deborah Elizabeth
Geburtsdatum	13. Januar 1974
Heimatort	Madiswil BE

Ausbildung

1987 – 1993	Literargymnasium Rämibühl
1993	Matura Typus B
1994 – 1999	Biologiestudium an der Eidgenössisch Technischen Hochschule Zürich Fachvertiefung in Ökologie und Systematik Fächer: Biosystematik, Gewässerwissenschaften, Meeresbiologie, Tierökologie und Umwelthygiene.
1998 – 1999	Diplomarbeit in Wildbiologie bei Prof. Dr. René Camenzind und Heinz Nigg Titel: Ermittlung des Lebensraumpotentials für Biber im Kanton Zürich unter Anwendung von GIS
2002 – 2009	Dissertation am Zoologischen Museum der Universität Zürich unter der Leitung von Prof. Dr. Georg Ribi Titel der Dissertation: Phylogeography, evolutionary history and genetic diversity of sea stars of the genus *Astropecten* and genetic structure within the Atlanto-Mediterranean species *A. aranciacus*